Computing the intrinsic mobility of electrons and holes in semiconductors

Part I : concepts and results

ABIDEV 2019

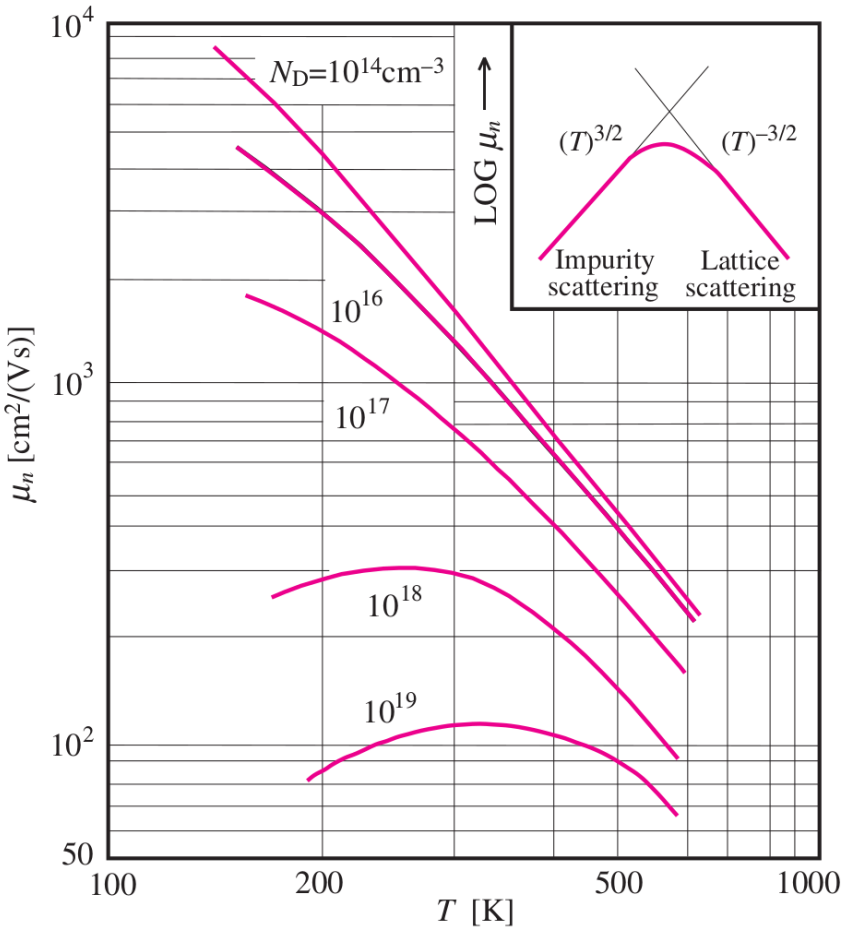
(1) Guillaume Brunin
(2) Henrique Miranda
Matteo Giantomassi
Gian-Marco Rignanese
Geoffroy Hautier

May 21, 2019

UCLouvain
Institute of Condensed Matter and Nanosciences

Electronic lifetimes and phonon-limited mobility

- Electron-phonon coupling dictates various phenomena : intrinsic mobility, indirect light absorption, superconductivity,...
- Electrons are scattered by phonons : mechanism dominating the electronic lifetimes and the mobility at high temperature



[Yu & Cardona, Fundamentals of Semiconductors]

Electronic lifetimes and phonon-limited mobility

Electron mobility in the relaxation-time approximation of the linearized Boltzmann transport formalism :

$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$

Electronic lifetimes and phonon-limited mobility

Electron mobility in the relaxation-time approximation of the linearized Boltzmann transport formalism :

$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$

Electronic lifetimes and phonon-limited mobility

Electron mobility in the relaxation-time approximation of the linearized Boltzmann transport formalism :

$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$

Electronic lifetimes due to the scattering by phonons :

$$\frac{1}{\tau_{n\mathbf{k}}} = \frac{2\pi}{\hbar} \sum_{m,\nu} \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 \times [(n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}})$$
$$+ (n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}})]$$

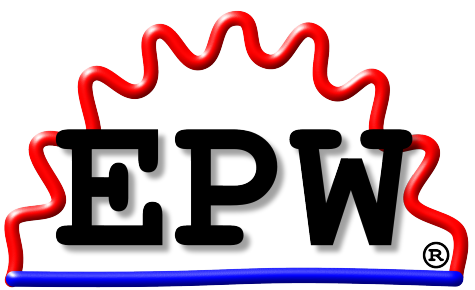
Electronic lifetimes and phonon-limited mobility

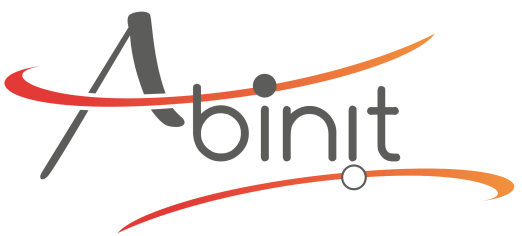
Electron mobility in the relaxation-time approximation of the linearized Boltzmann transport formalism :

$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$

Electronic lifetimes due to the scattering by phonons :

$$\frac{1}{\tau_{n\mathbf{k}}} = \frac{2\pi}{\hbar} \sum_{m,\nu} \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 \times [(n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}})$$
$$+ (n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}})]$$





PERTURBO

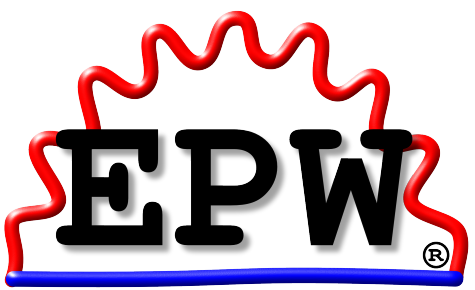
Electronic lifetimes and phonon-limited mobility

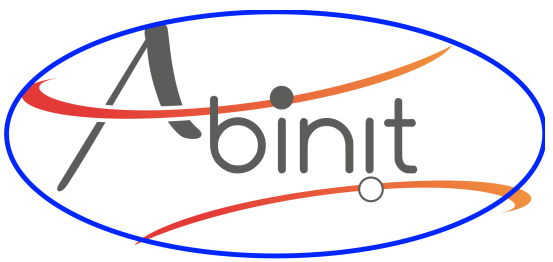
Electron mobility in the relaxation-time approximation of the linearized Boltzmann transport formalism :

$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$

Electronic lifetimes due to the scattering by phonons :

$$\frac{1}{\tau_{n\mathbf{k}}} = \frac{2\pi}{\hbar} \sum_{m,\nu} \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 \times [(n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}})$$
$$+ (n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}})]$$



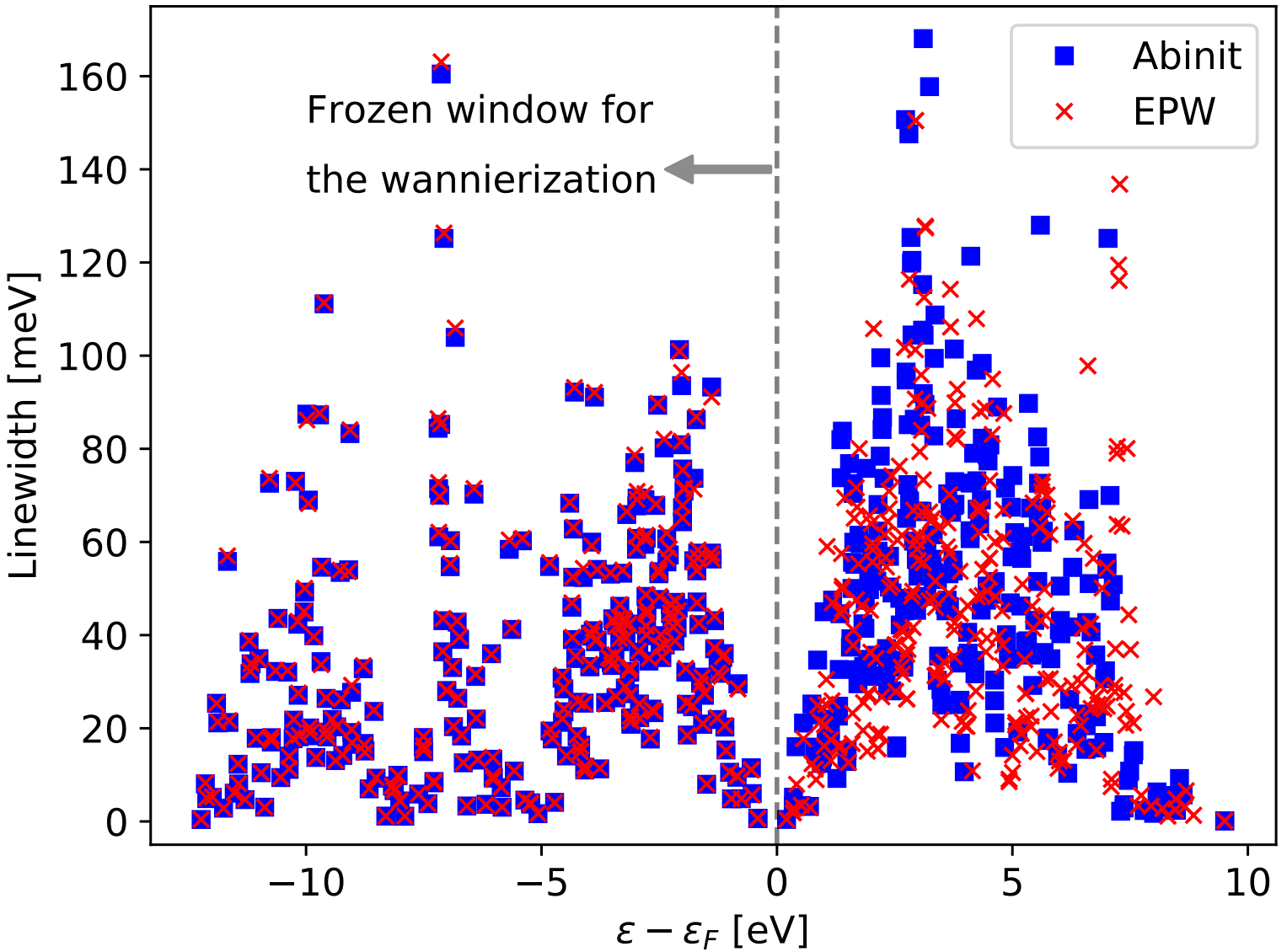


PERTURBO

No need for Wannier functions or atomic orbitals !

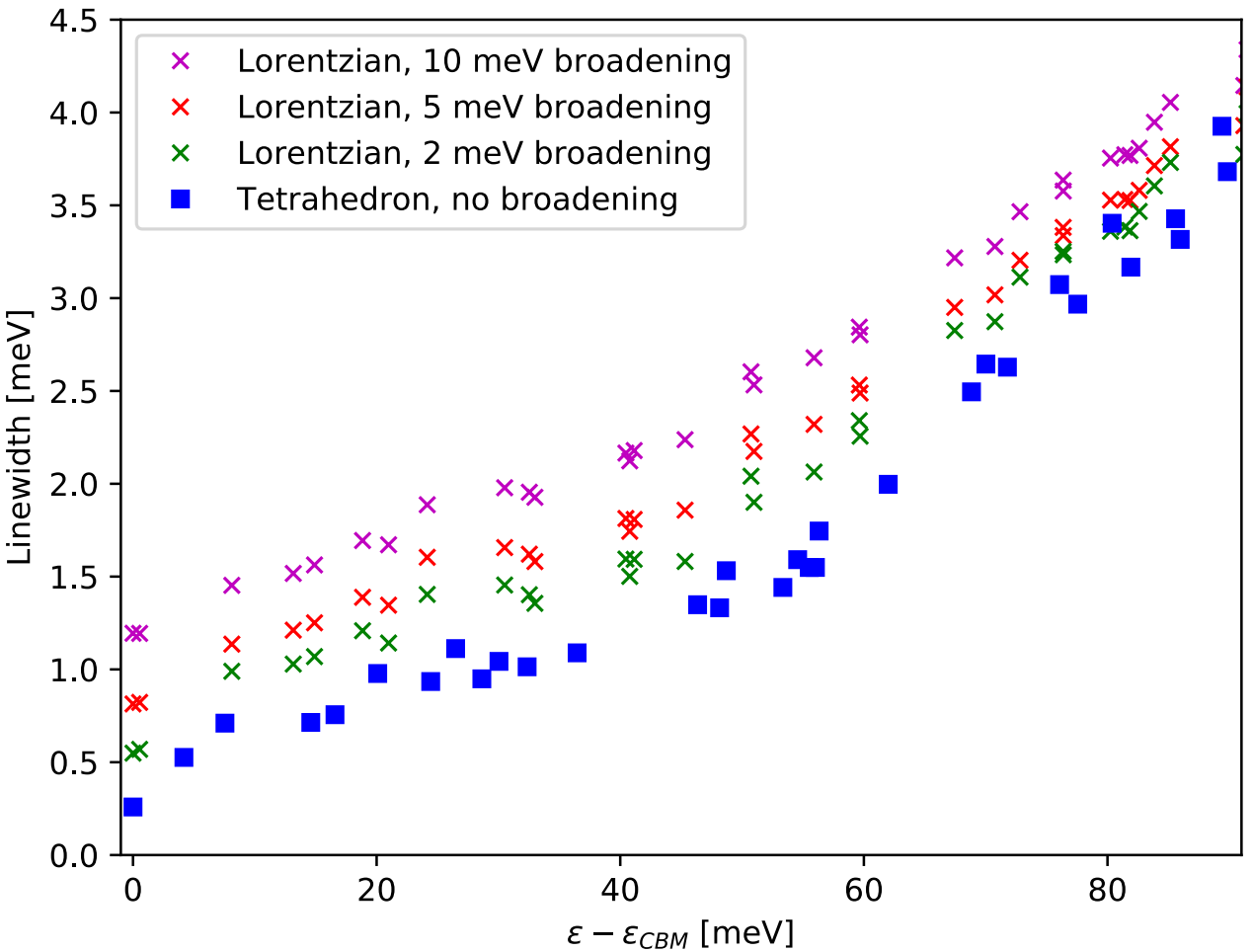
Lifetimes in Silicon : agreement between ABINIT and EPW

- Lifetime $= 1/\text{linewidth}$



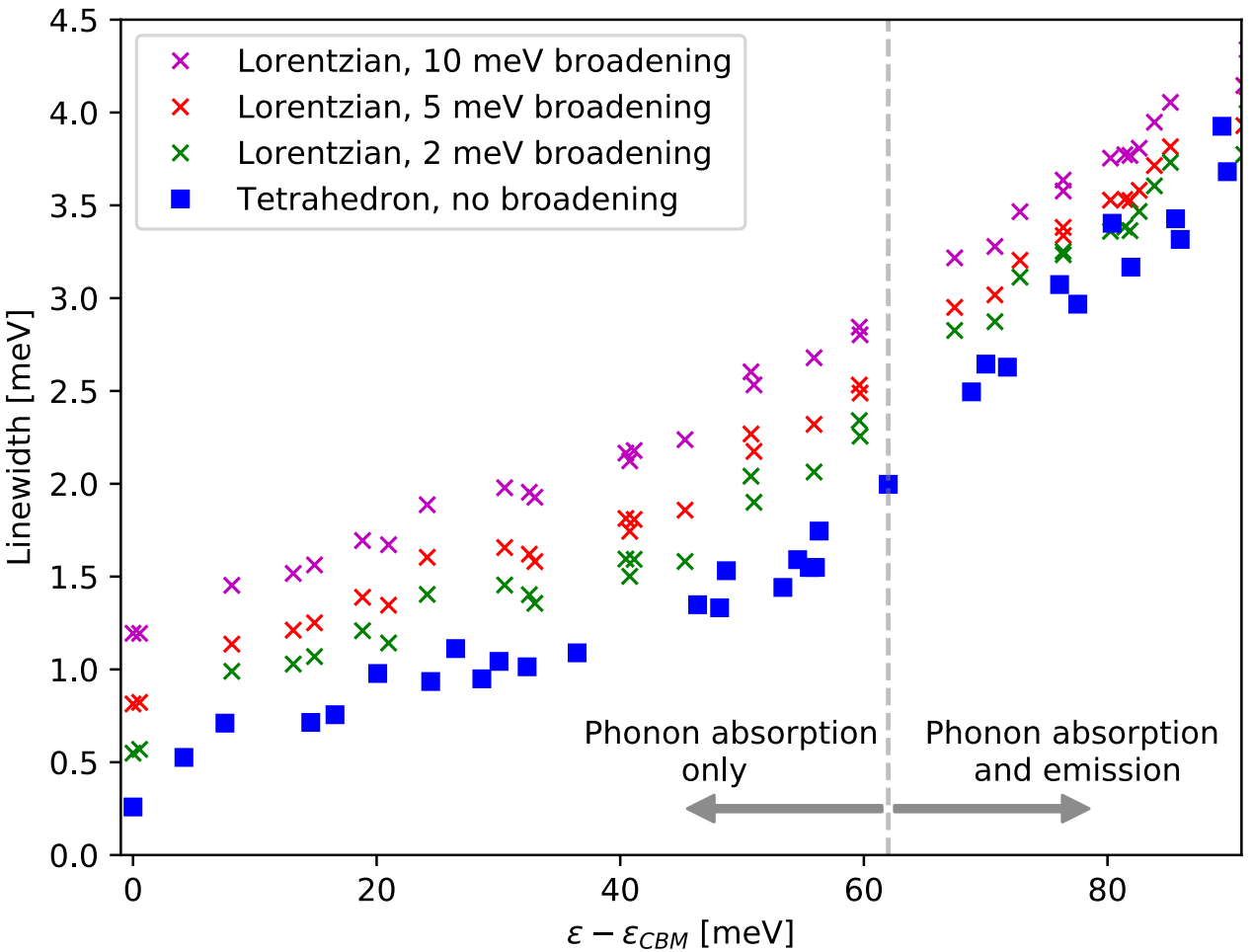
Tetrahedron integration to replace Lorentzian broadening

- Linewidth $\propto \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} \pm \omega_{\nu\mathbf{q}})$
- Linewidth around the CBM of Si for a $60 \times 60 \times 60$ **k**-point grid (300K)



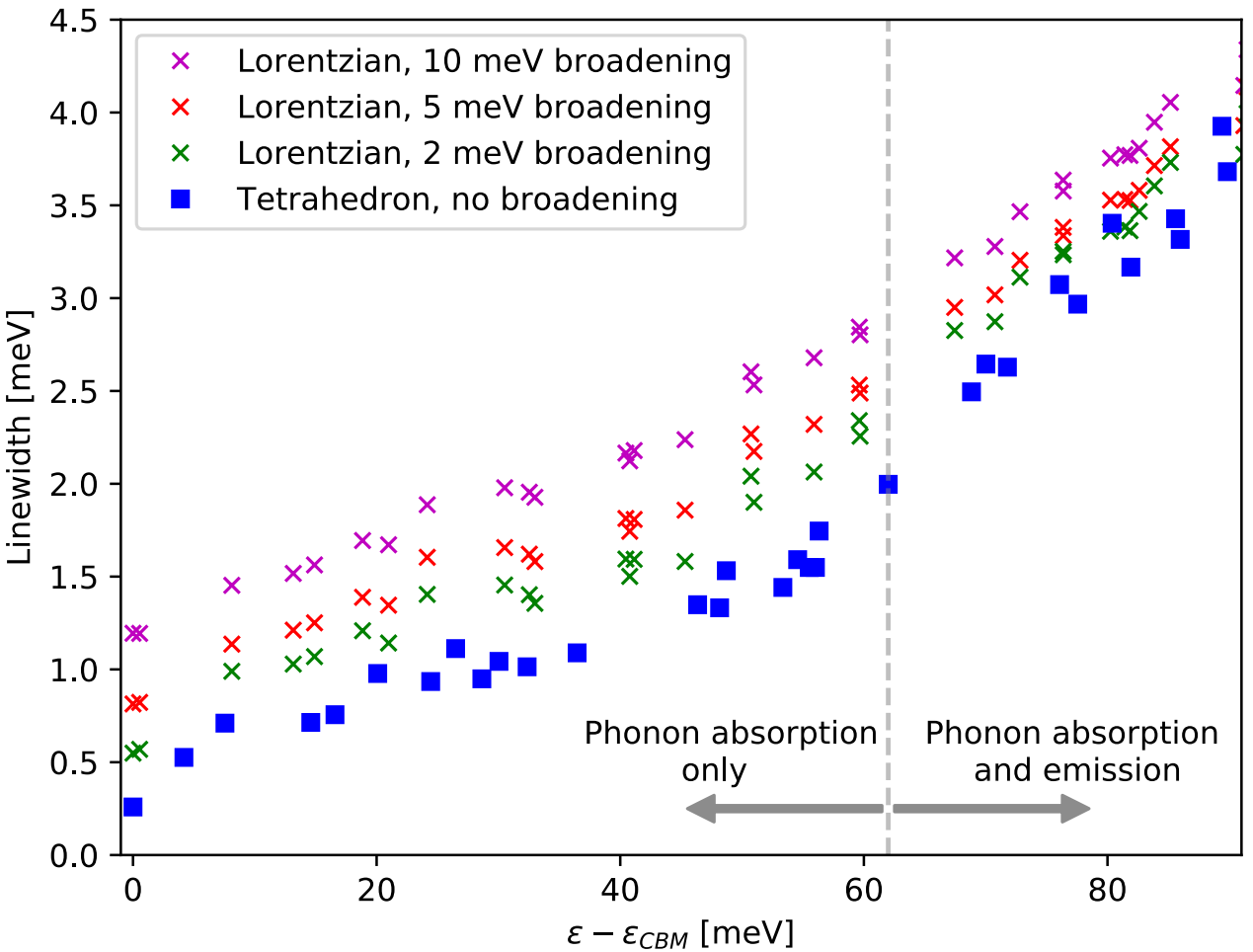
Tetrahedron integration to replace Lorentzian broadening

- Linewidth $\propto \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} \pm \omega_{\nu\mathbf{q}})$
- Linewidth around the CBM of Si for a $60 \times 60 \times 60$ **k**-point grid (300K)
- More scattering channels for $\varepsilon - \varepsilon_{CBM} > \omega_{LO} \Rightarrow$ larger linewidths



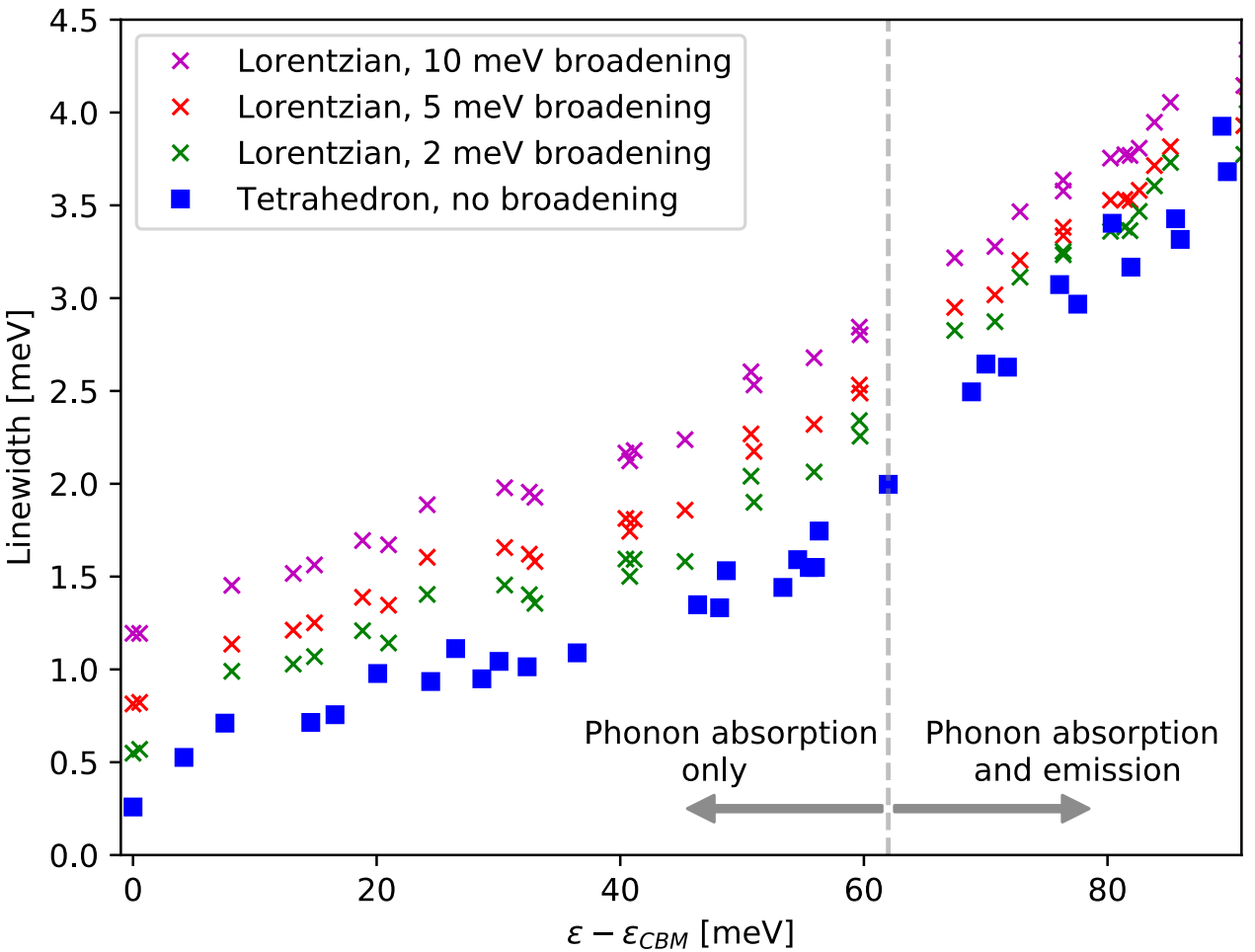
Tetrahedron integration to replace Lorentzian broadening

- Linewidth $\propto \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} \pm \omega_{\nu\mathbf{q}})$
- Linewidth around the CBM of Si for a $60 \times 60 \times 60$ **k**-point grid (300K)
- More scattering channels for $\varepsilon - \varepsilon_{CBM} > \omega_{LO} \Rightarrow$ larger linewidths
- A Lorentzian broadening requires convergence studies



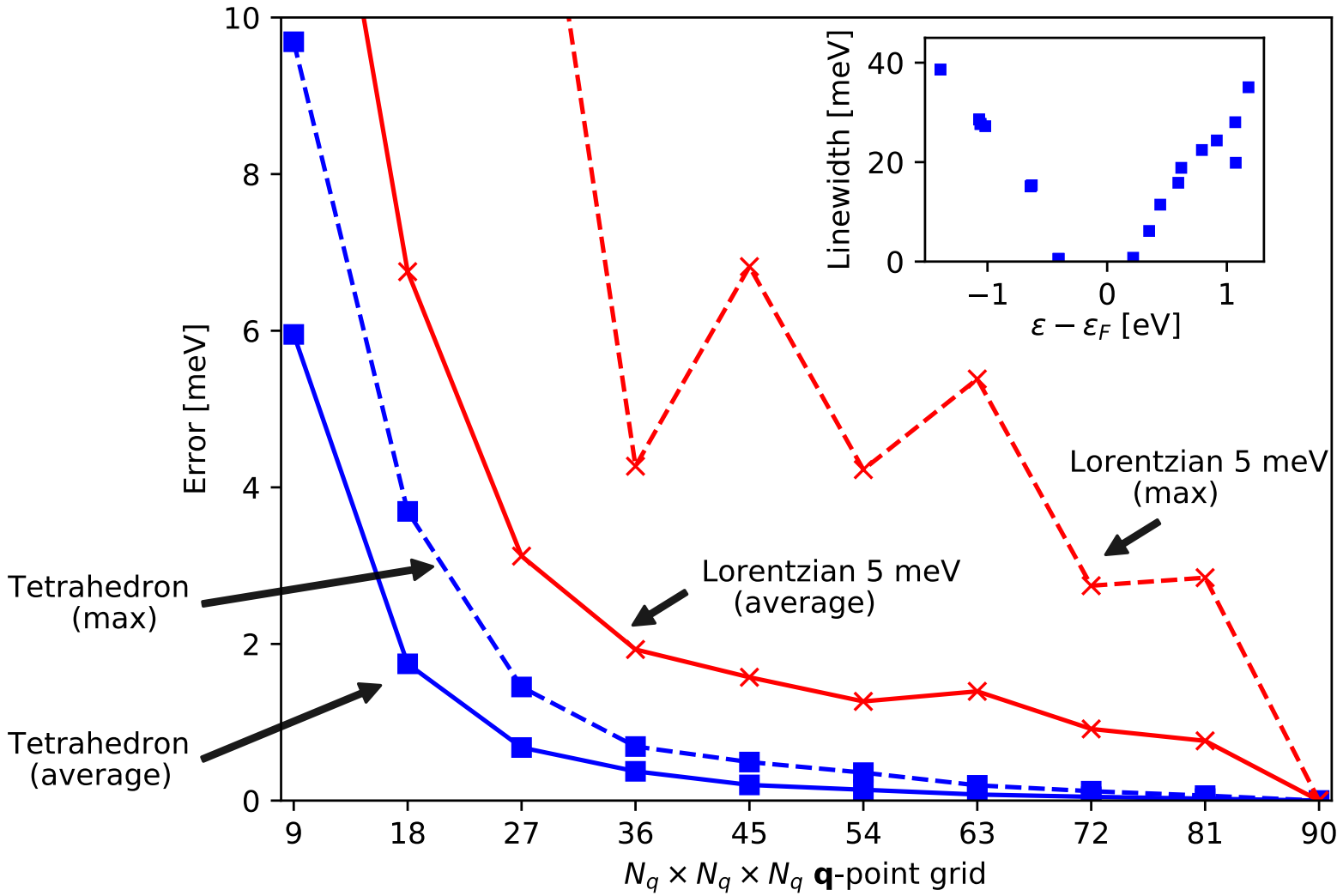
Tetrahedron integration to replace Lorentzian broadening

- Linewidth $\propto \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} \pm \omega_{\nu\mathbf{q}})$
- Linewidth around the CBM of Si for a $60 \times 60 \times 60$ **k**-point grid (300K)
- More scattering channels for $\varepsilon - \varepsilon_{CBM} > \omega_{LO} \Rightarrow$ larger linewidths
- A Lorentzian broadening requires convergence studies
- Tetrahedron method : no broadening parameter



Convergence of the linewidths in Silicon

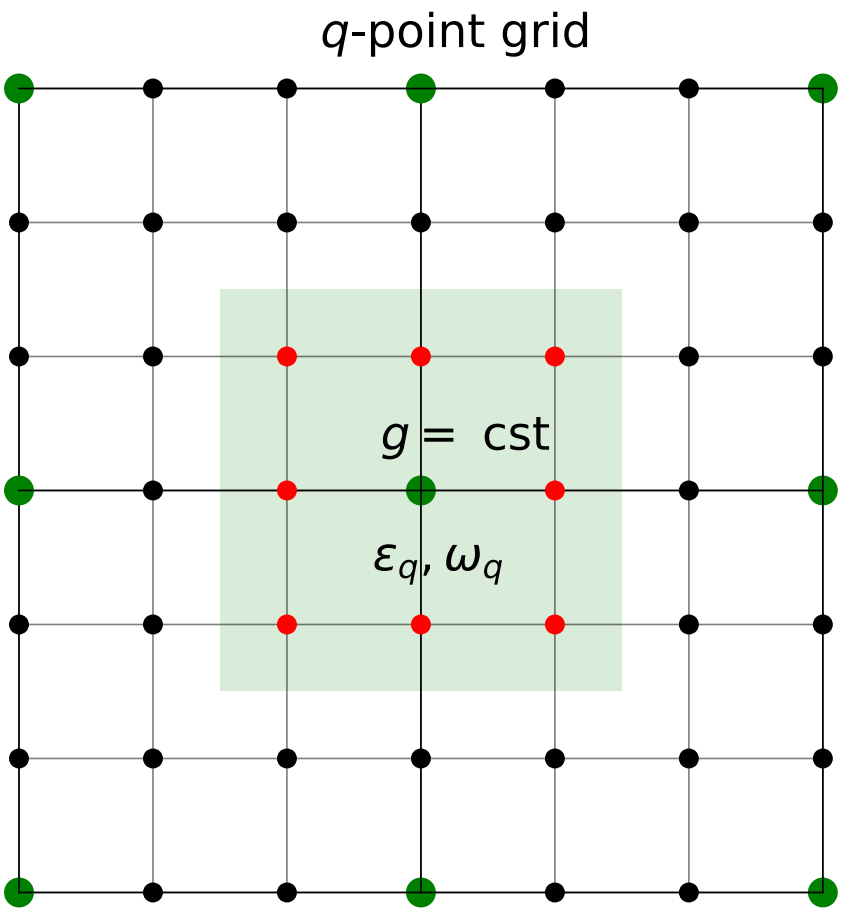
- Linewidth $\propto \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 \ldots$
- Linewidths on a $9 \times 9 \times 9$ **k**-point grid, for increasing **q**-point grids
- The Tetrahedron integration converges fast (like for a DOS !)



Double-grid technique : 1 grid for matrix elements, 1 grid for energies

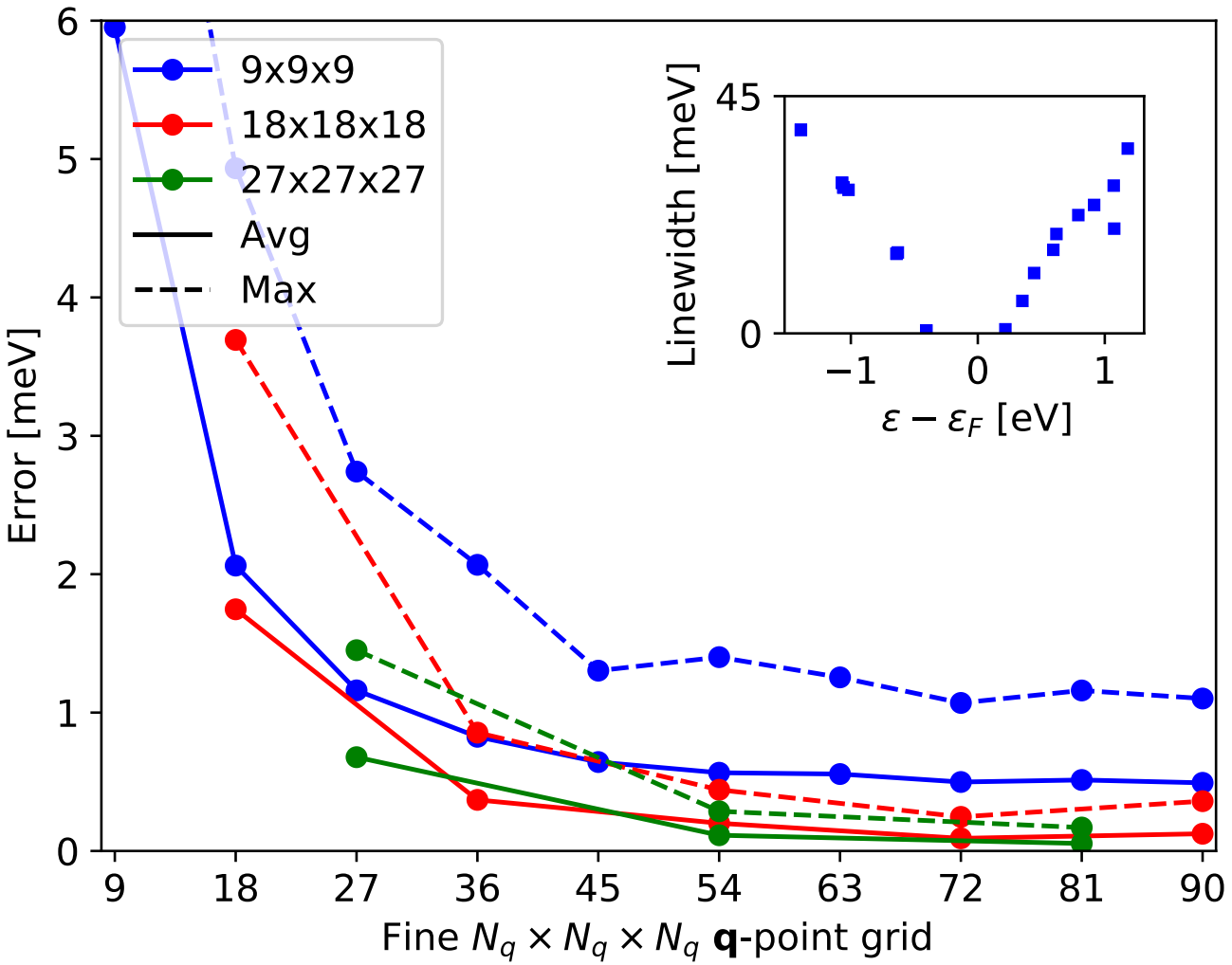
$$\frac{1}{\tau_{n\mathbf{k}}} = \frac{2\pi}{\hbar} \sum_{m,\nu} \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 \times [(n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}})\delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}})$$

$$+(n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}})\delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}})]$$



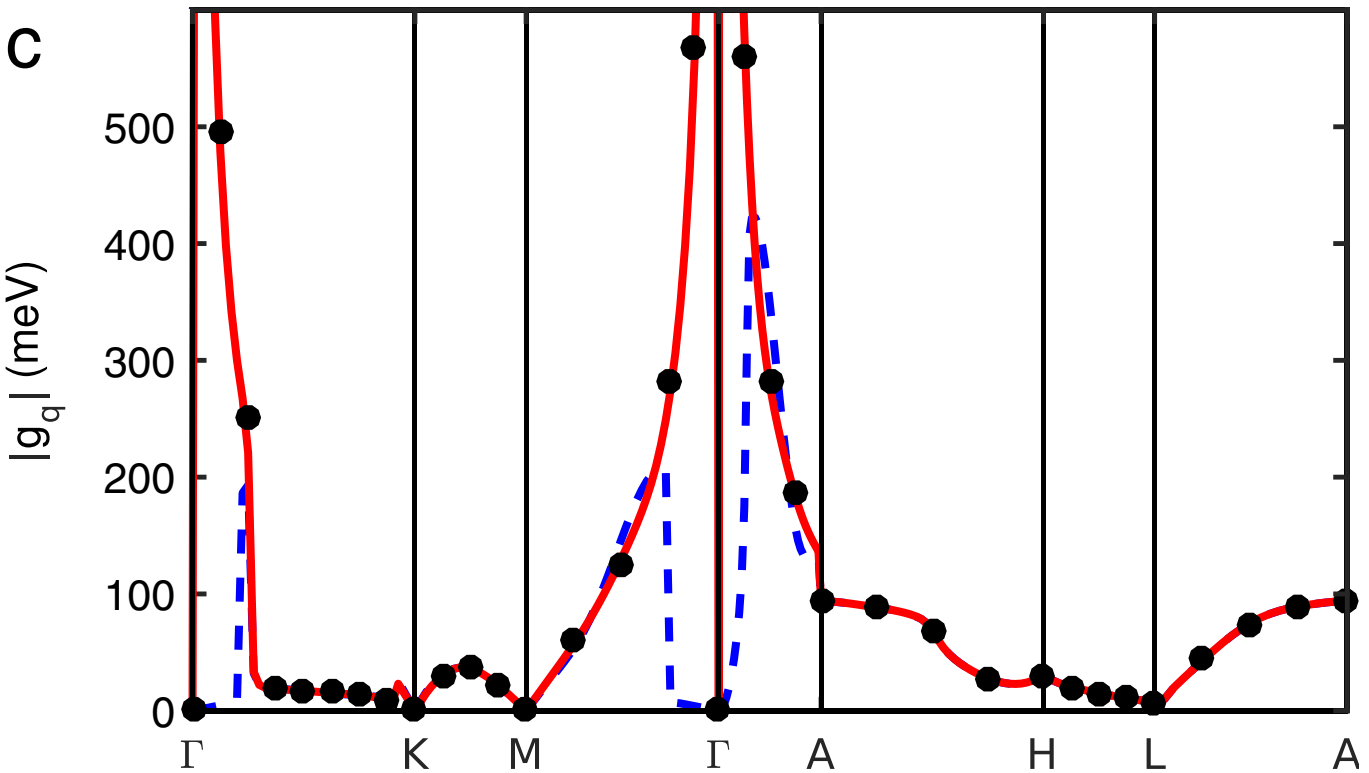
Double-grid technique : errors on the linewidths in Silicon

- Linewidth $\propto \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 ...$
- Linewidths on a $9 \times 9 \times 9$ **k**-point grid
- Blue : $9 \times 9 \times 9$ **q**-point grid for matrix elements, increasing density for the energy grid only



Polar materials : divergence of the matrix elements

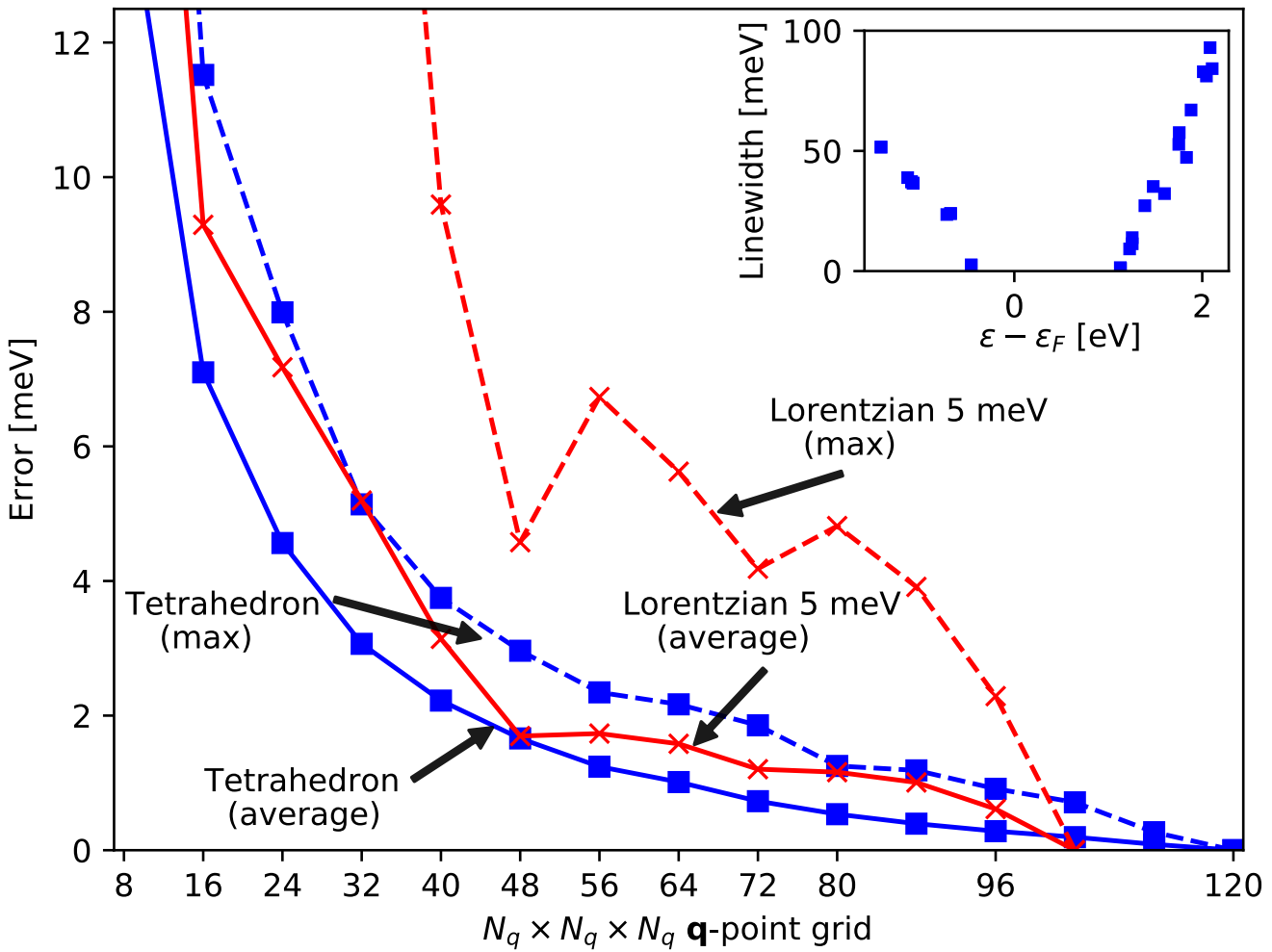
$$g^{LR}_{mn,\nu}(\mathbf{k},\mathbf{q}) = i\frac{4\pi}{\Omega}\frac{e^2}{4\pi\varepsilon_0}\sum_{\kappa}\left(\frac{\hbar}{2NM_\kappa\omega_{\nu\mathbf{q}}}\right)^{(1/2)}\sum_{\mathbf{G}\neq-\mathbf{q}}\frac{(\mathbf{q}+\mathbf{G})\cdot\mathbf{Z}^*_\kappa\cdot\mathbf{e}_{\kappa\nu}(\mathbf{q})}{(\mathbf{q}+\mathbf{G})\cdot\varepsilon^\infty\cdot(\mathbf{q}+\mathbf{G})}$$
$$\times\langle\Psi_{m\mathbf{k}+\mathbf{q}}|e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}|\Psi_{n\mathbf{k}}\rangle$$



[Poncé, CPC 209, 116-133 (2016)]

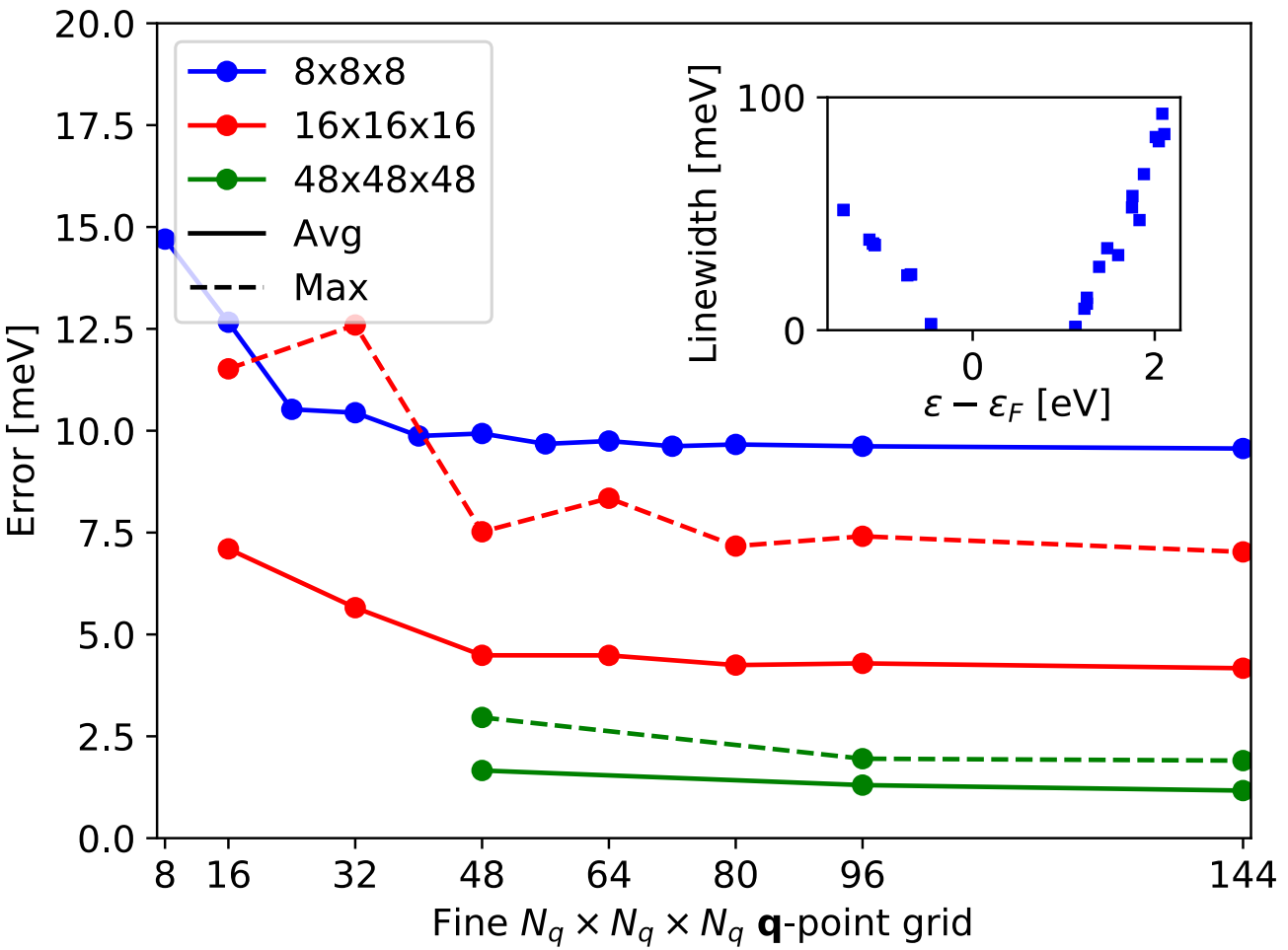
Convergence of the linewidths in GaP

- Linewidth $\propto \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 \ldots$
- Linewidths on a $8 \times 8 \times 8$ **k**-point grid, for increasing **q**-point grids
- The Tetrahedron integration converges slower than in Silicon
- This is due to the polar divergence of electron-phonon matrix elements



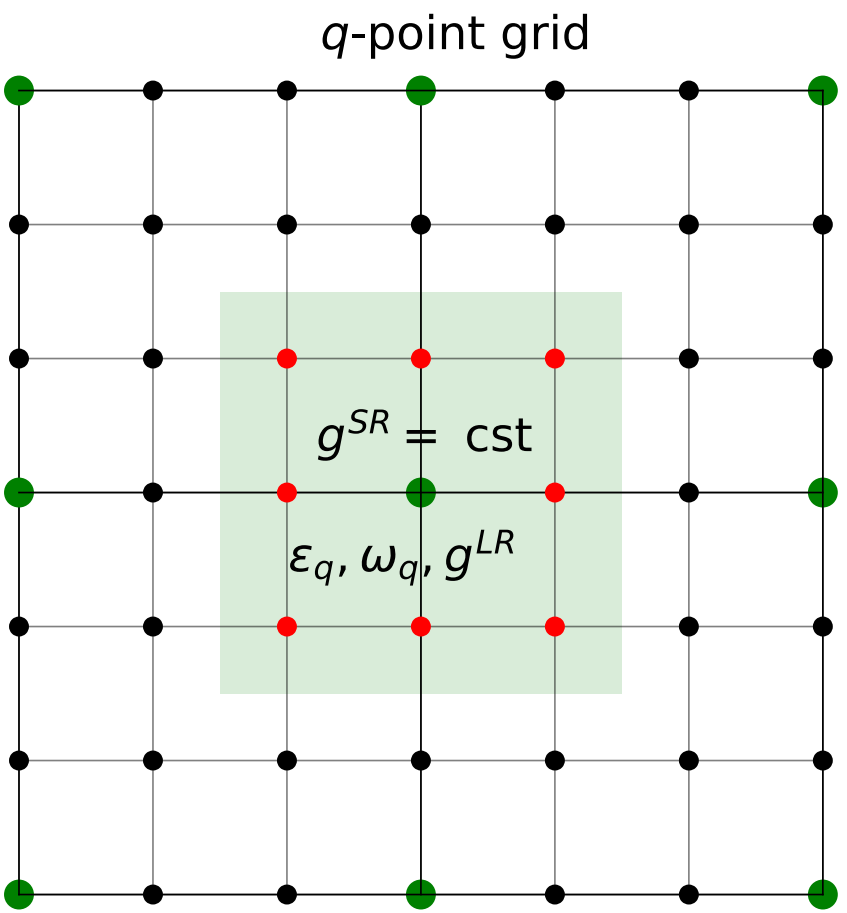
Double-grid technique for polar materials : less efficient...

- Linewidth $\propto \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2$...
- Linewidths on a $8 \times 8 \times 8$ **k**-point grid
- Blue : $8 \times 8 \times 8$ **q**-point grid for matrix elements, increasing density for the energy grid only



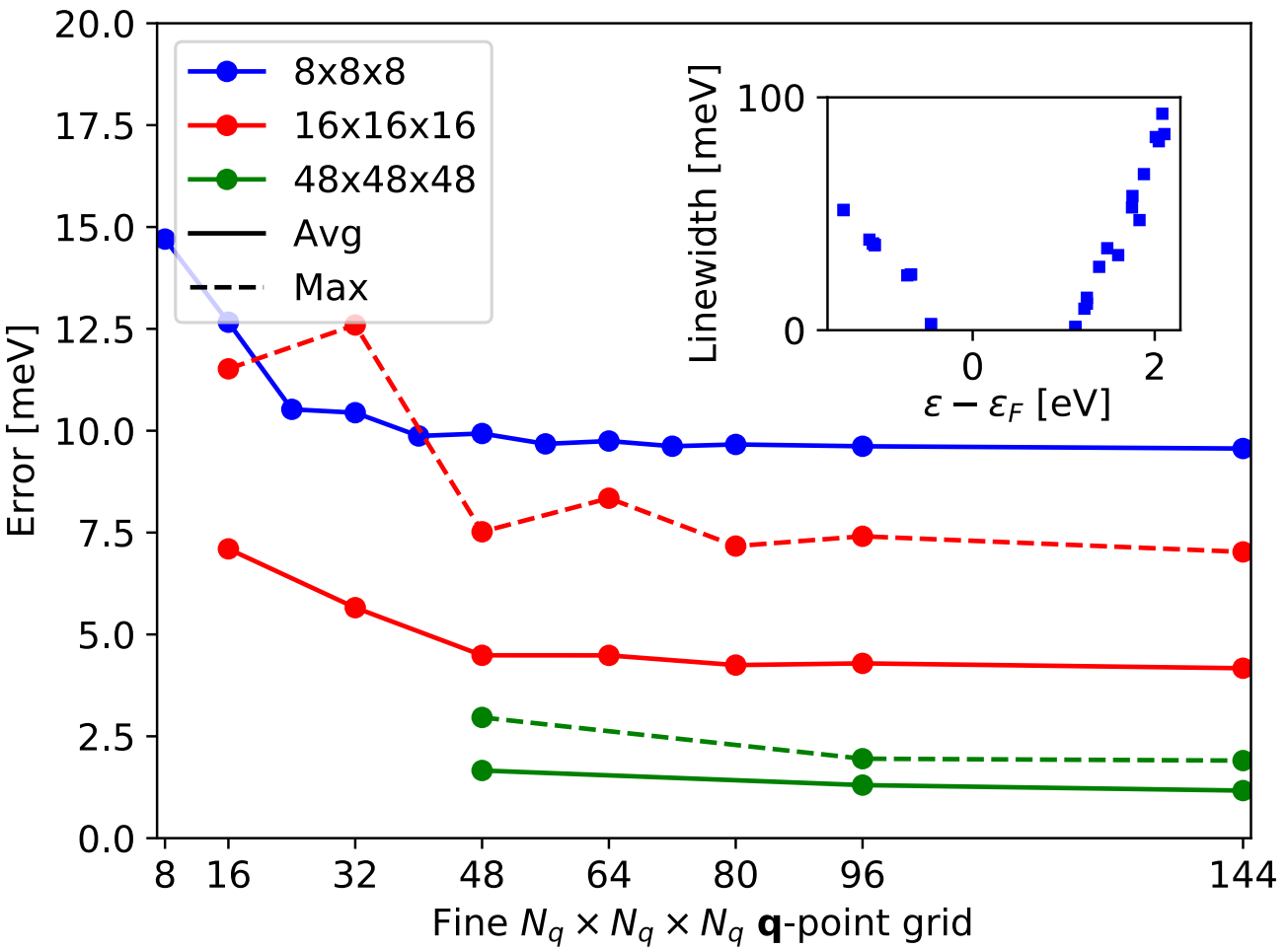
Double-grid technique for polar materials : special treatment

$$\frac{1}{\tau_{n\mathbf{k}}} = \frac{2\pi}{\hbar} \sum_{m,\nu} \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |\underbrace{g_{mn,\nu}(\mathbf{k},\mathbf{q})}_{g^{LR}+g^{SR}}|^2 \times [(n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}})\delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}})$$
$$+(n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}})\delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}})]$$



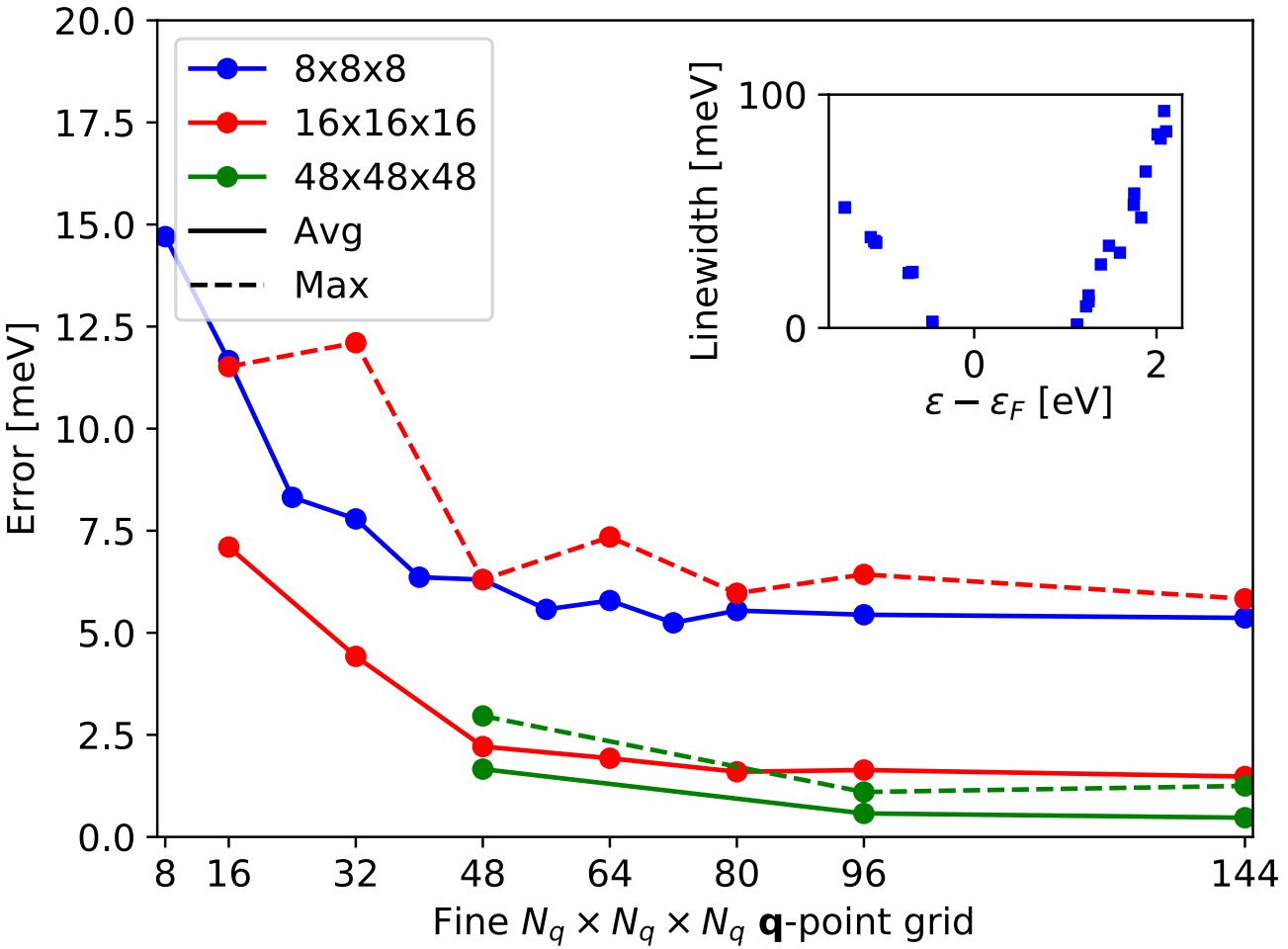
Double-grid technique for polar materials : faster convergence

- Linewidth $\propto \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 ...$
- Linewidths on a $8 \times 8 \times 8$ **k**-point grid
- Blue : $8 \times 8 \times 8$ **q**-point grid for matrix elements, increasing density for the energy grid only



Double-grid technique for polar materials : faster convergence

- Linewidth $\propto \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 ...$
- Linewidths on a $8 \times 8 \times 8$ **k**-point grid
- Blue : $8 \times 8 \times 8$ **q**-point grid for matrix elements, increasing density for the energy and g^{LR} only



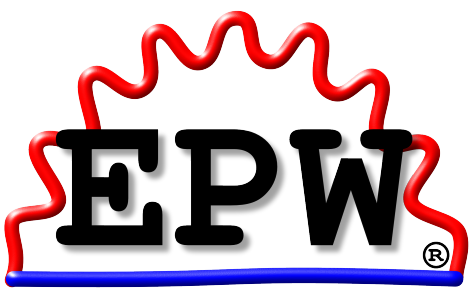
Electronic lifetimes and phonon-limited mobility

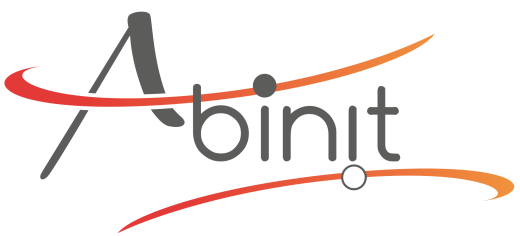
Electron mobility in the relaxation-time approximation of the linearized Boltzmann transport formalism :

$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$

Electronic lifetimes due to the scattering by phonons :

$$\frac{1}{\tau_{n\mathbf{k}}} = \frac{2\pi}{\hbar} \sum_{m,\nu} \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 \times [(n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}})$$
$$+ (n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}}) \delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}})]$$



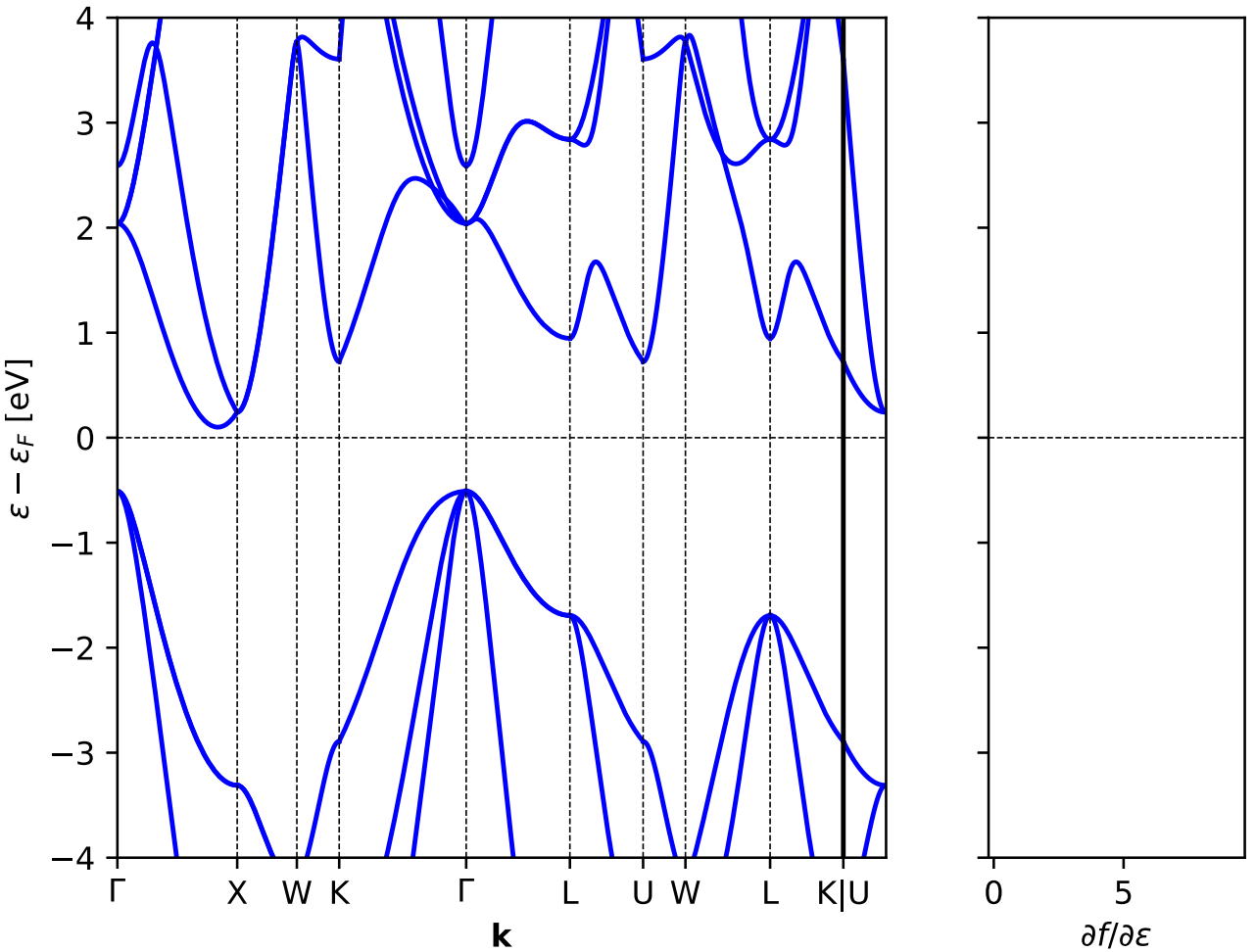


PERTURBO

Transport properties : do we *really* need the lifetimes everywhere ?

The mobility contains a derivative of the Fermi-Dirac occupation function :

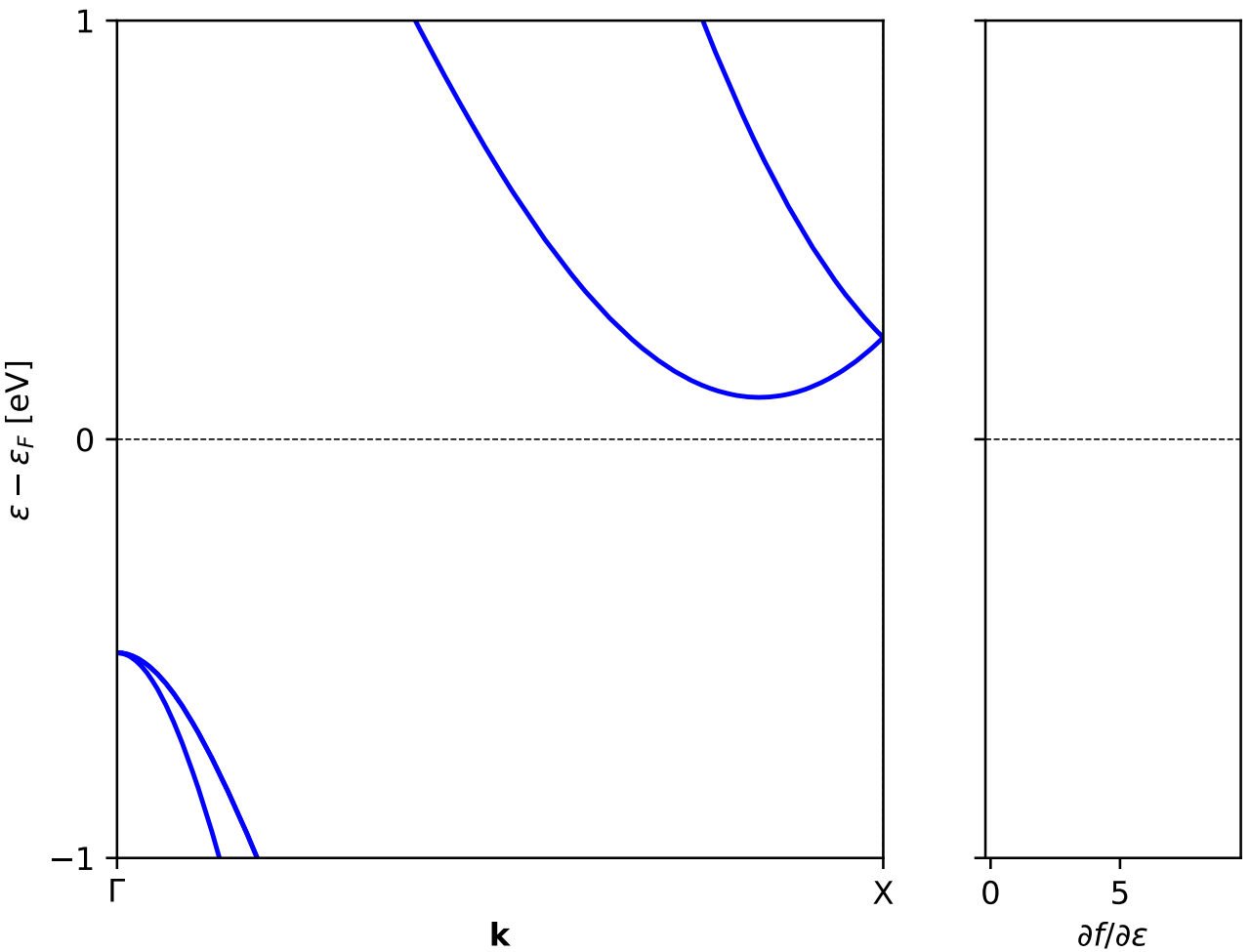
$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$



Transport properties : do we *really* need the lifetimes everywhere ?

The mobility contains a derivative of the Fermi-Dirac occupation function :

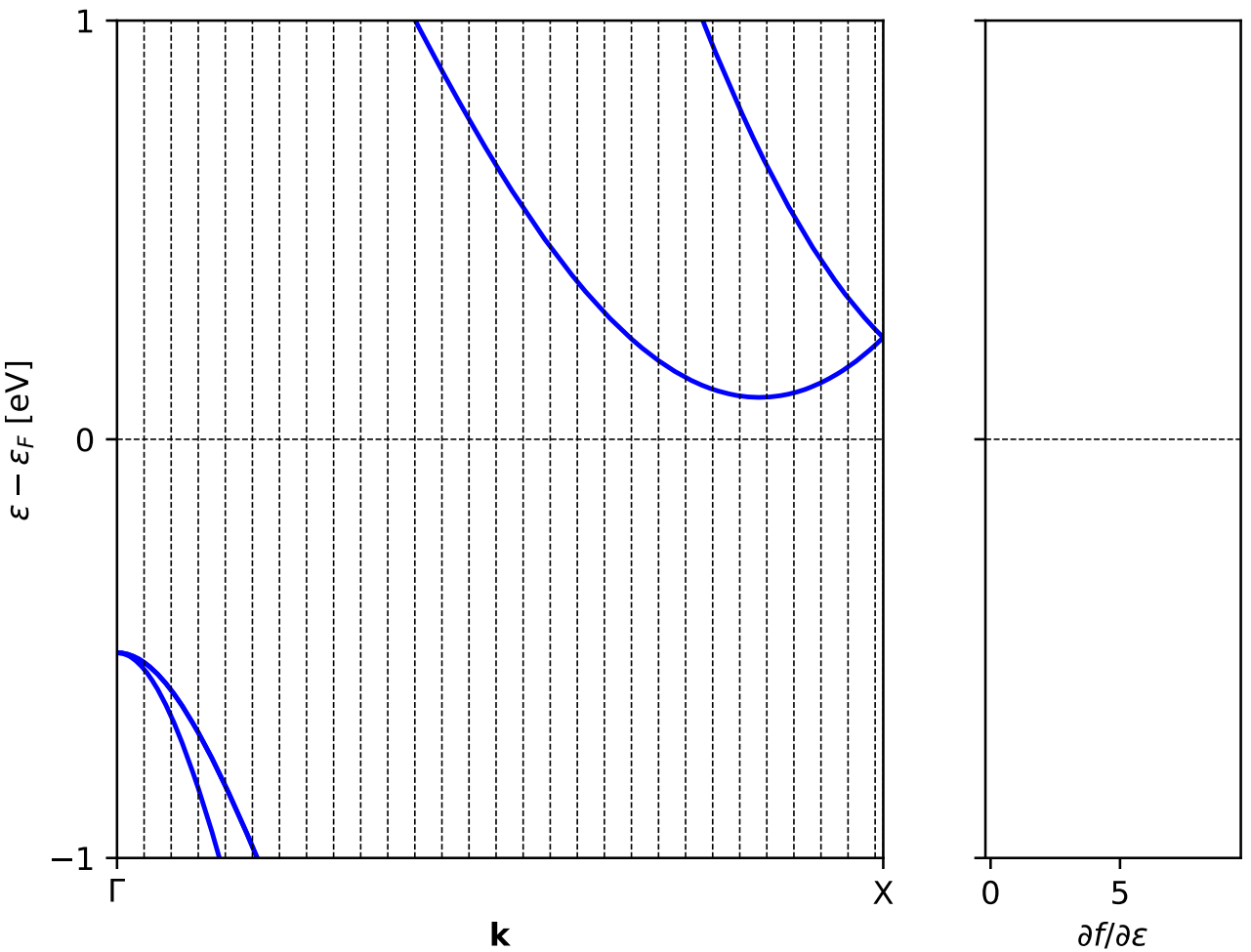
$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$



Transport properties : do we *really* need the lifetimes everywhere ?

The mobility contains a derivative of the Fermi-Dirac occupation function :

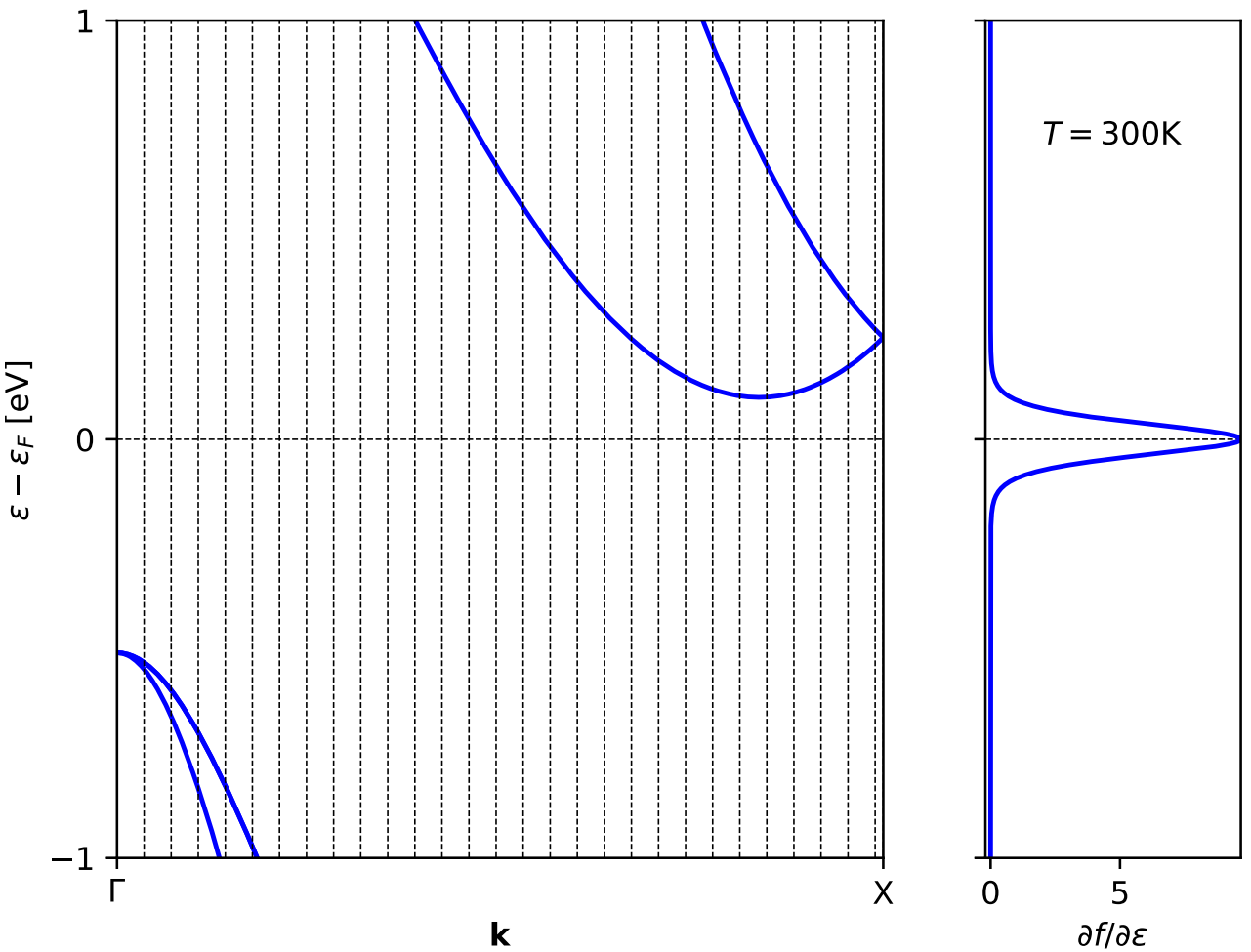
$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$



Transport properties : do we *really* need the lifetimes everywhere ?

The mobility contains a derivative of the Fermi-Dirac occupation function :

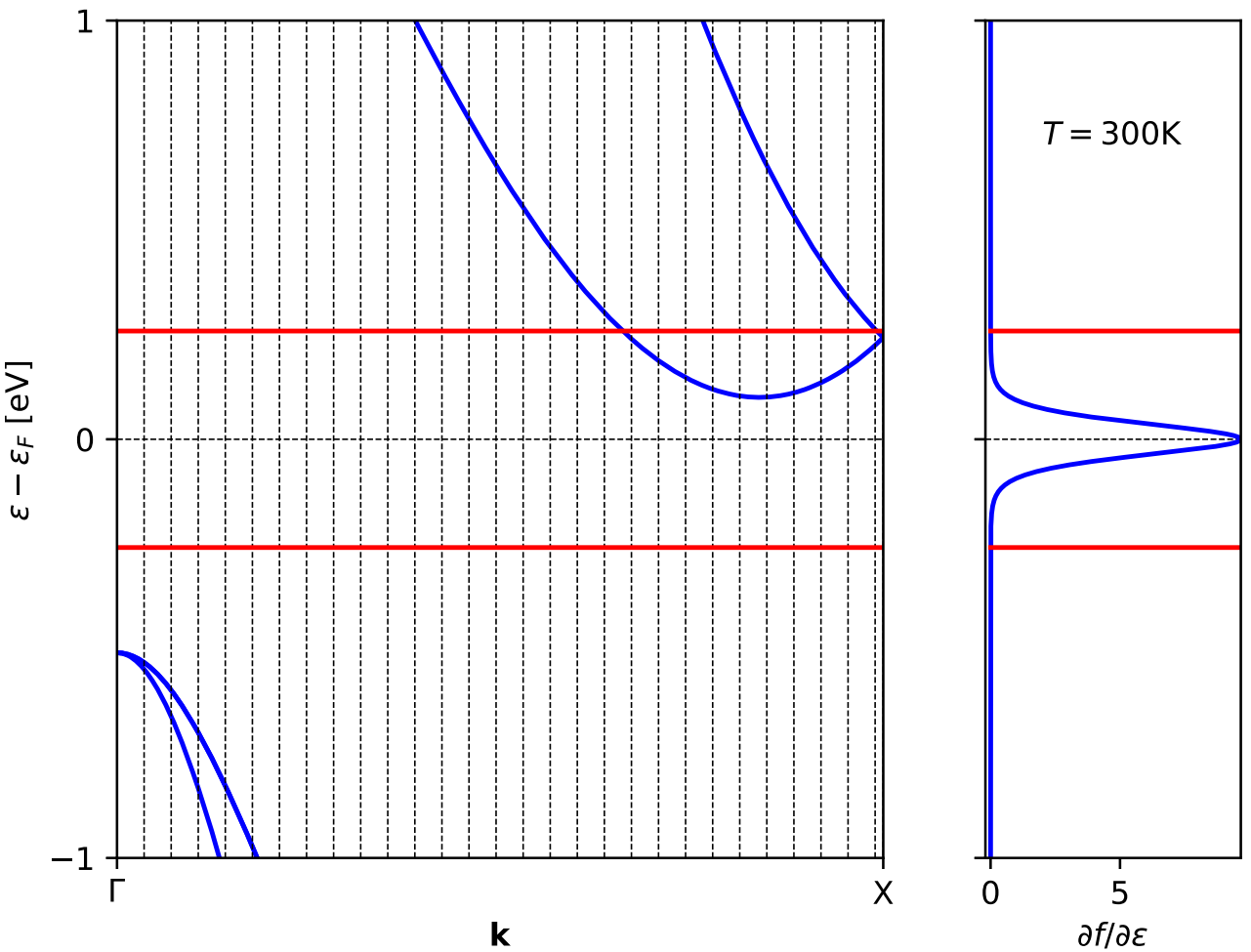
$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$



Transport properties : do we *really* need the lifetimes everywhere ?

The mobility contains a derivative of the Fermi-Dirac occupation function :

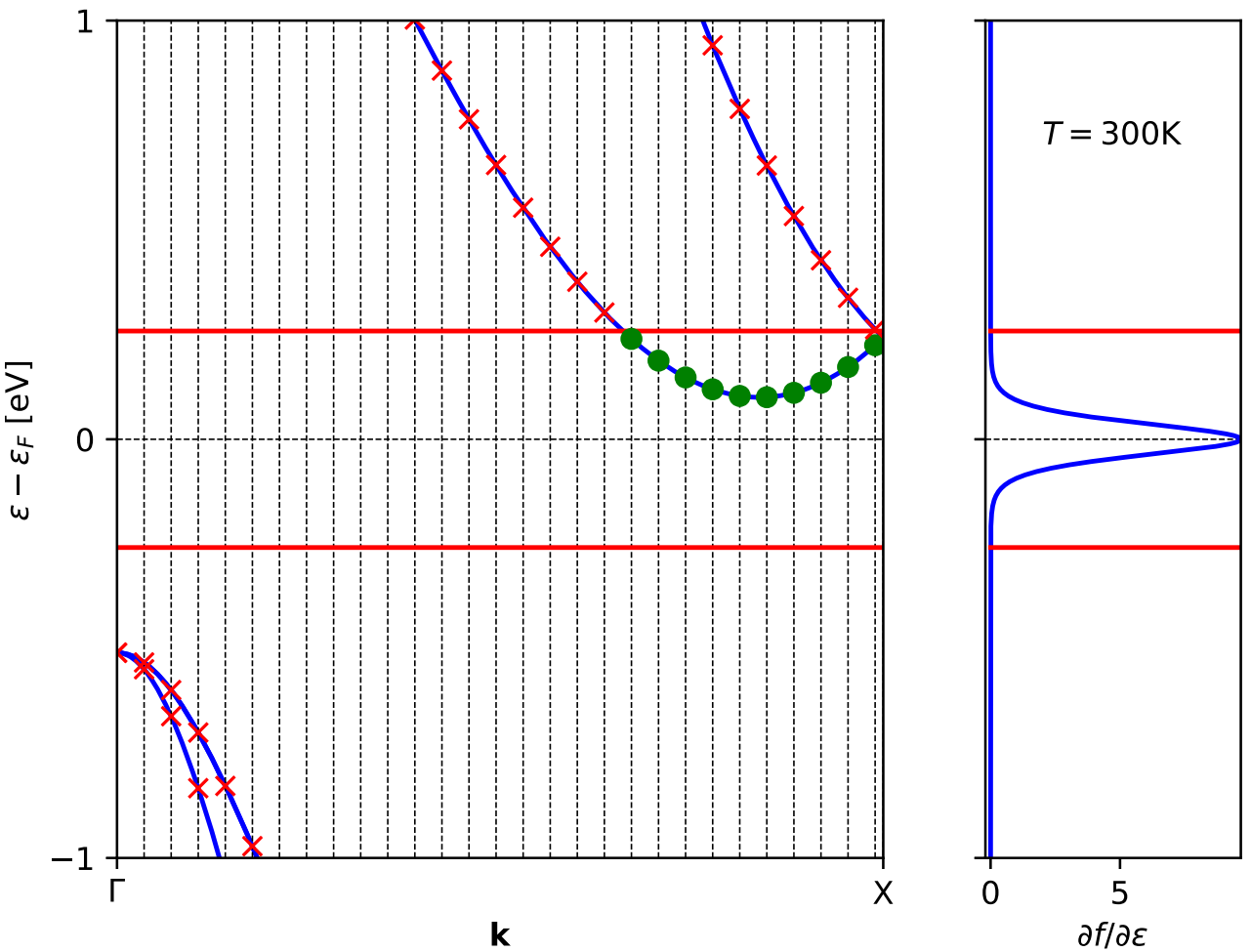
$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$



Transport properties : do we *really* need the lifetimes everywhere ?

The mobility contains a derivative of the Fermi-Dirac occupation function :

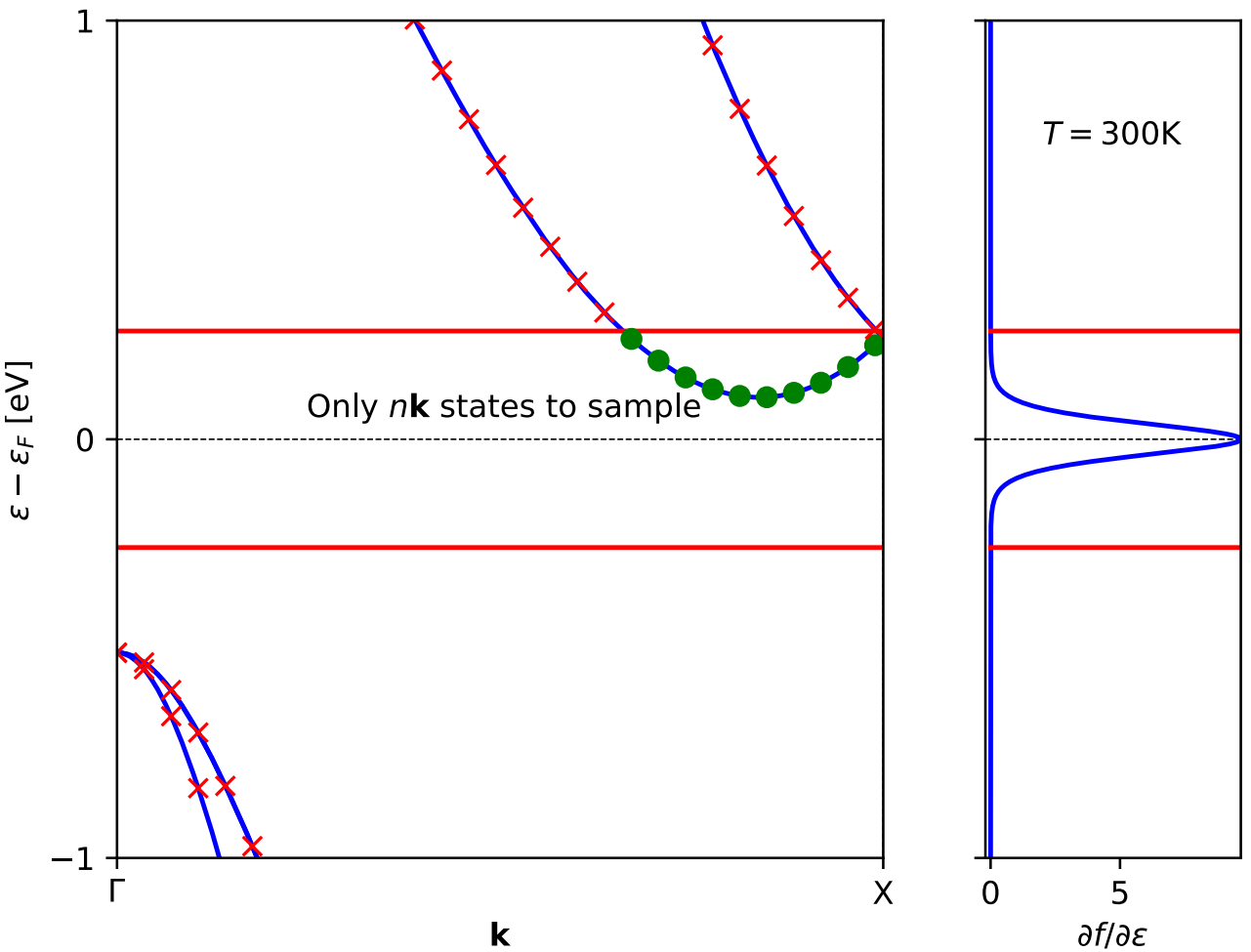
$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$



Transport properties : do we *really* need the lifetimes everywhere ?

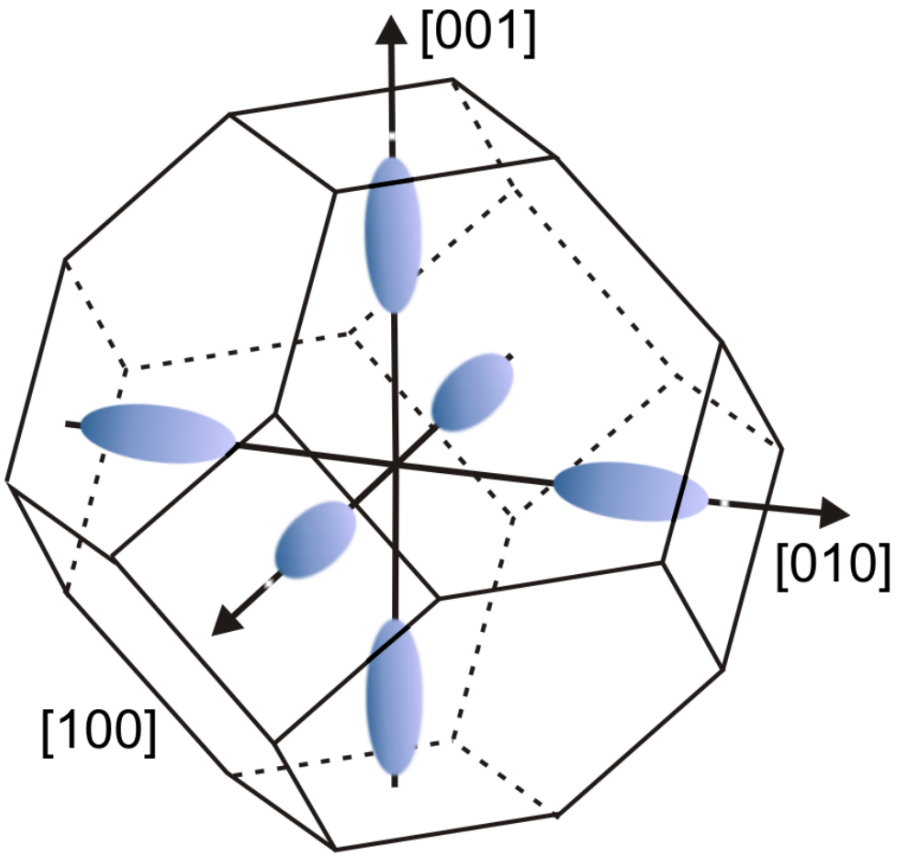
The mobility contains a derivative of the Fermi-Dirac occupation function :

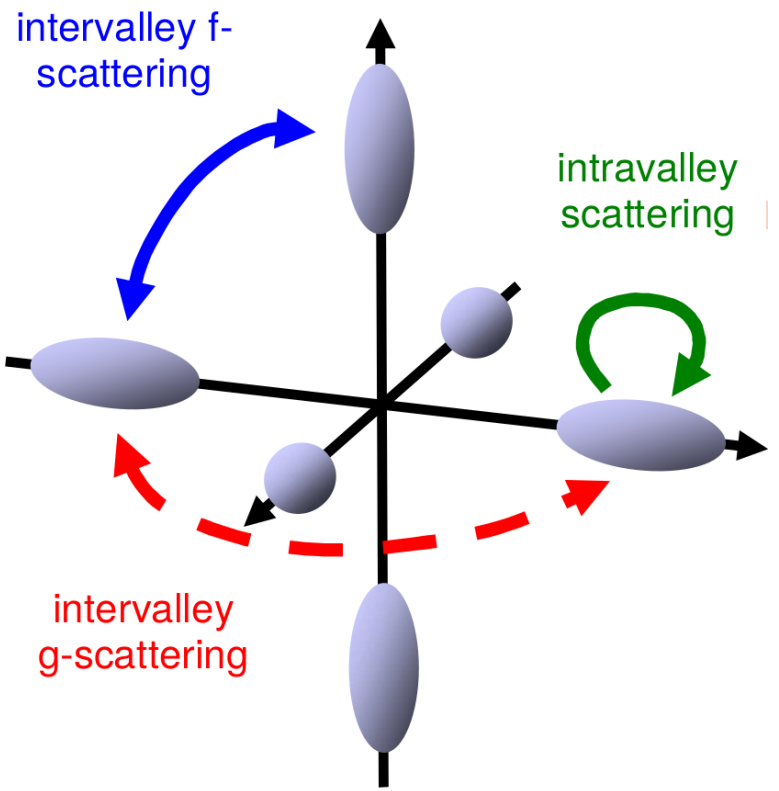
$$\mu_{e,\alpha\beta}(\varepsilon_F, T) = \frac{-e}{\Omega n_e} \sum_{n \in CB} \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon}$$



Lifetimes in CB pockets : do we *really* need to sample all **q**-points ?

- Transitions from **k** to **k** + **q** : momentum conservation
- **q**-points are also limited by the energy conservation : $\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} = \pm\omega_{\nu\mathbf{q}}$
- Only a limited set of **q**-points contribute and need to be taken into account !





Mobility of electrons in Silicon with ABINIT

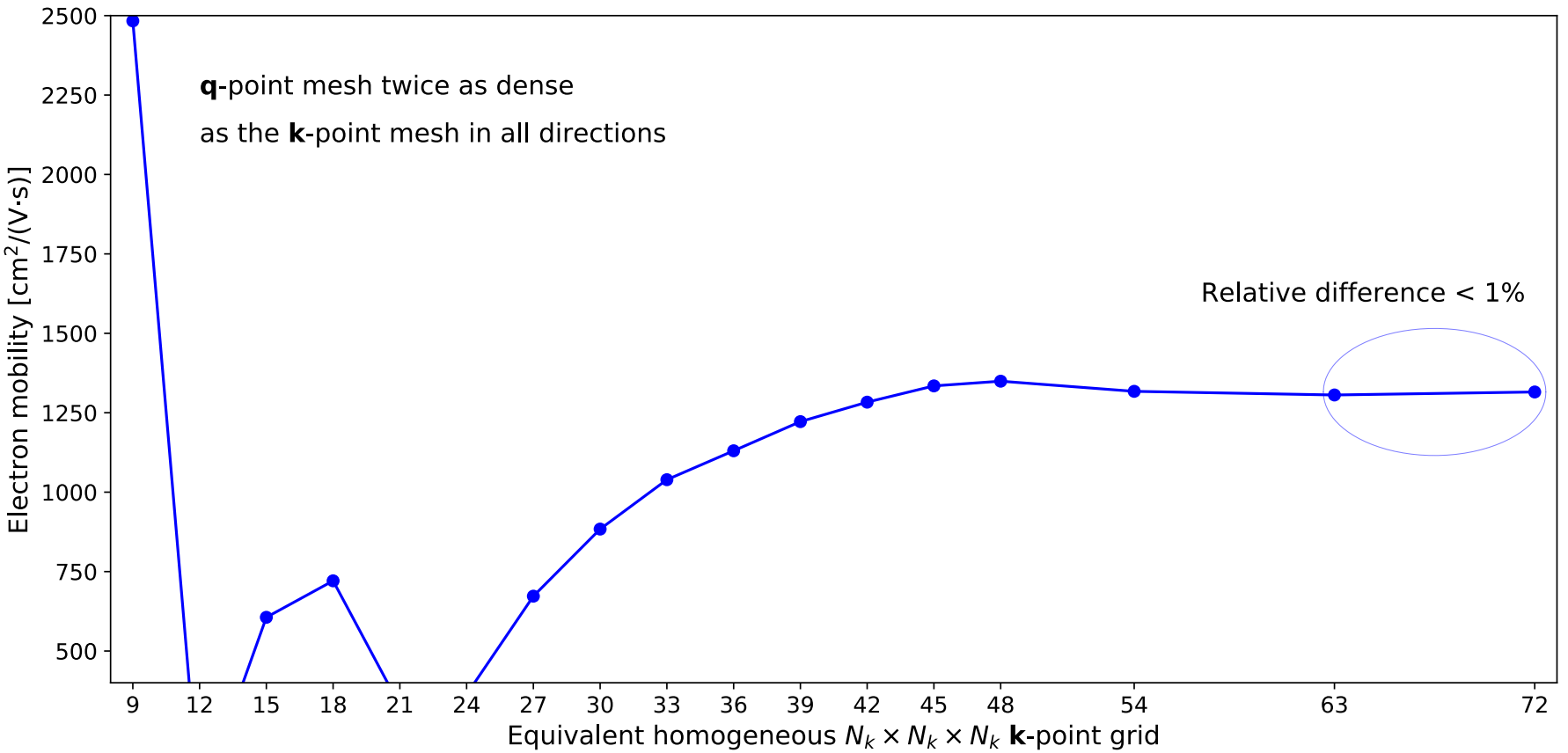
New implementation in ABINIT :

- Selects the important **k**- and **q**-points
- Computes the lifetimes and velocities for these points
- Performs the integration to obtain the phonon-limited mobility

Mobility of electrons in Silicon with ABINIT

New implementation in ABINIT :

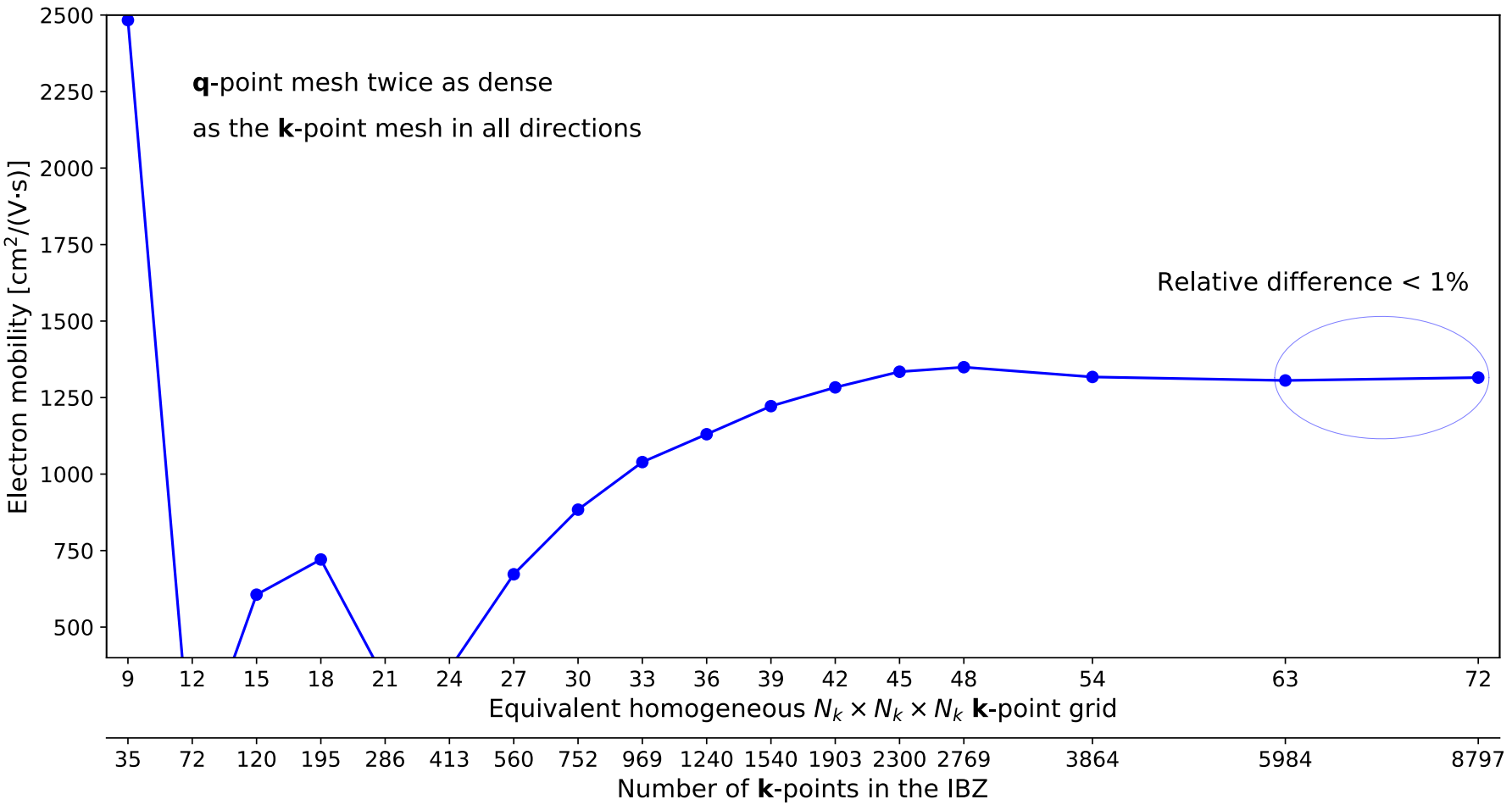
- Selects the important **k**- and **q**-points
- Computes the lifetimes and velocities for these points
- Performs the integration to obtain the phonon-limited mobility



Mobility of electrons in Silicon with ABINIT

New implementation in ABINIT :

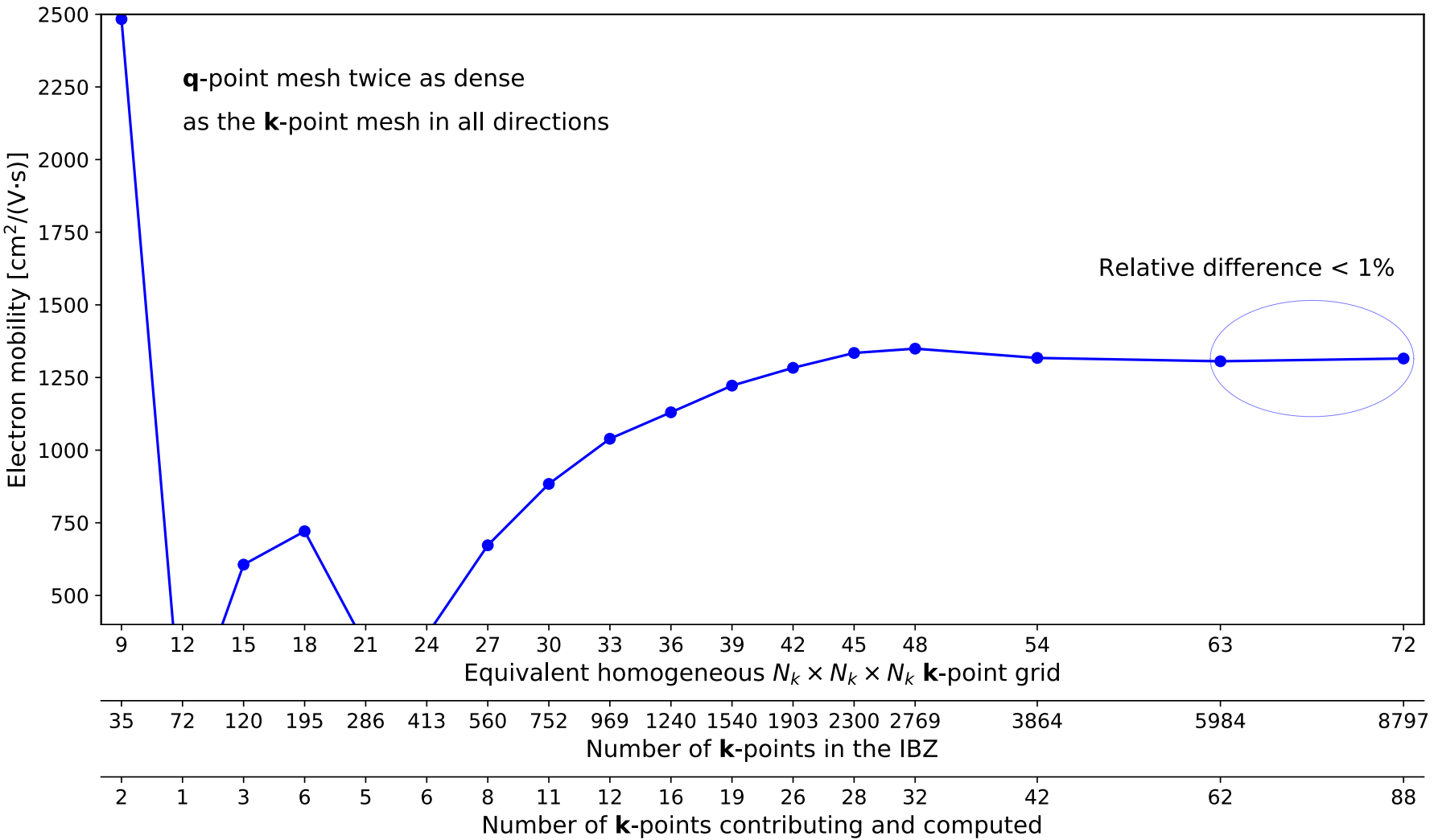
- Selects the important **k**- and **q**-points
- Computes the lifetimes and velocities for these points
- Performs the integration to obtain the phonon-limited mobility



Mobility of electrons in Silicon with ABINIT

New implementation in ABINIT :

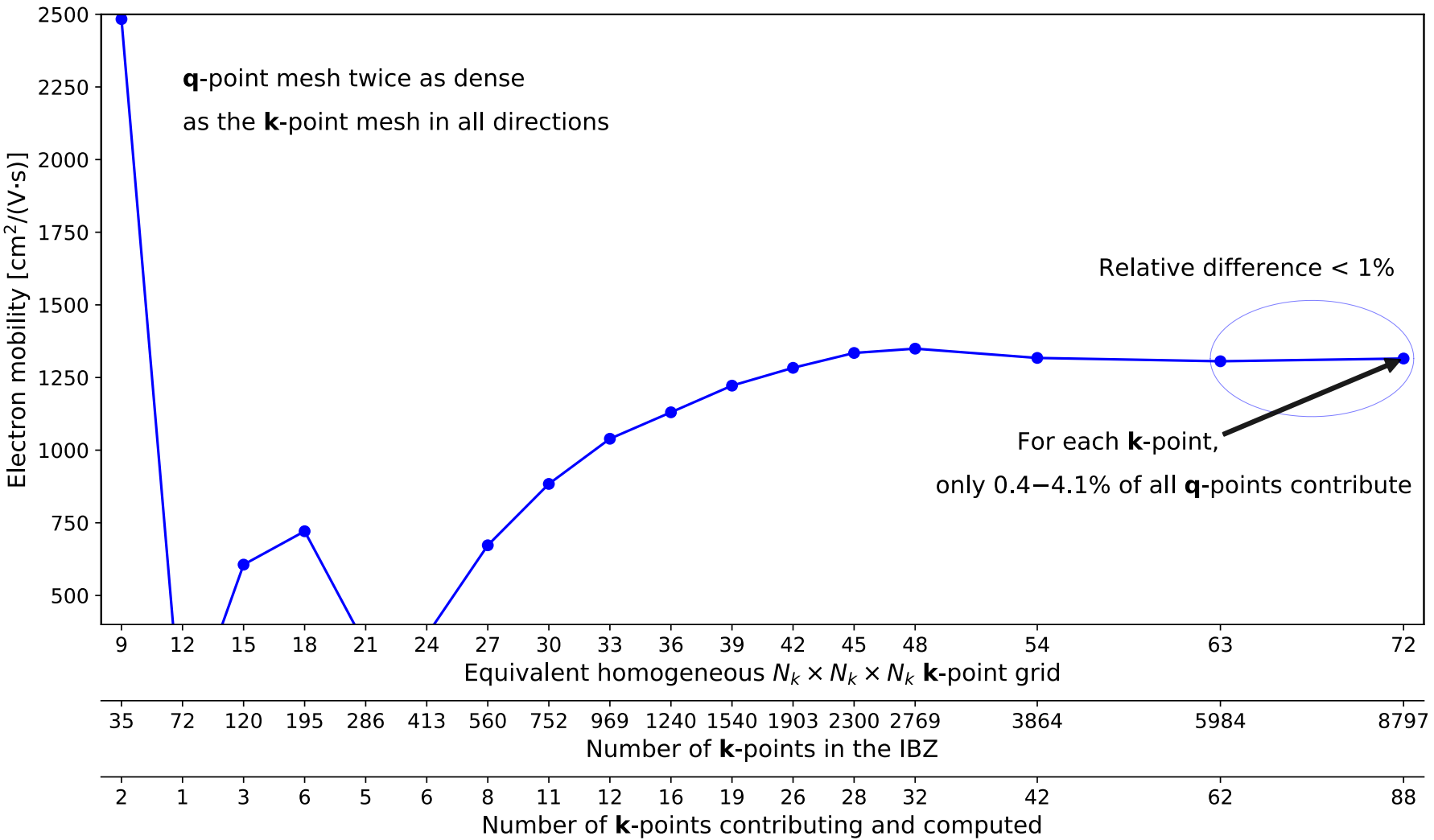
- Selects the important **k**- and **q**-points
- Computes the lifetimes and velocities for these points
- Performs the integration to obtain the phonon-limited mobility



Mobility of electrons in Silicon with ABINIT

New implementation in ABINIT :

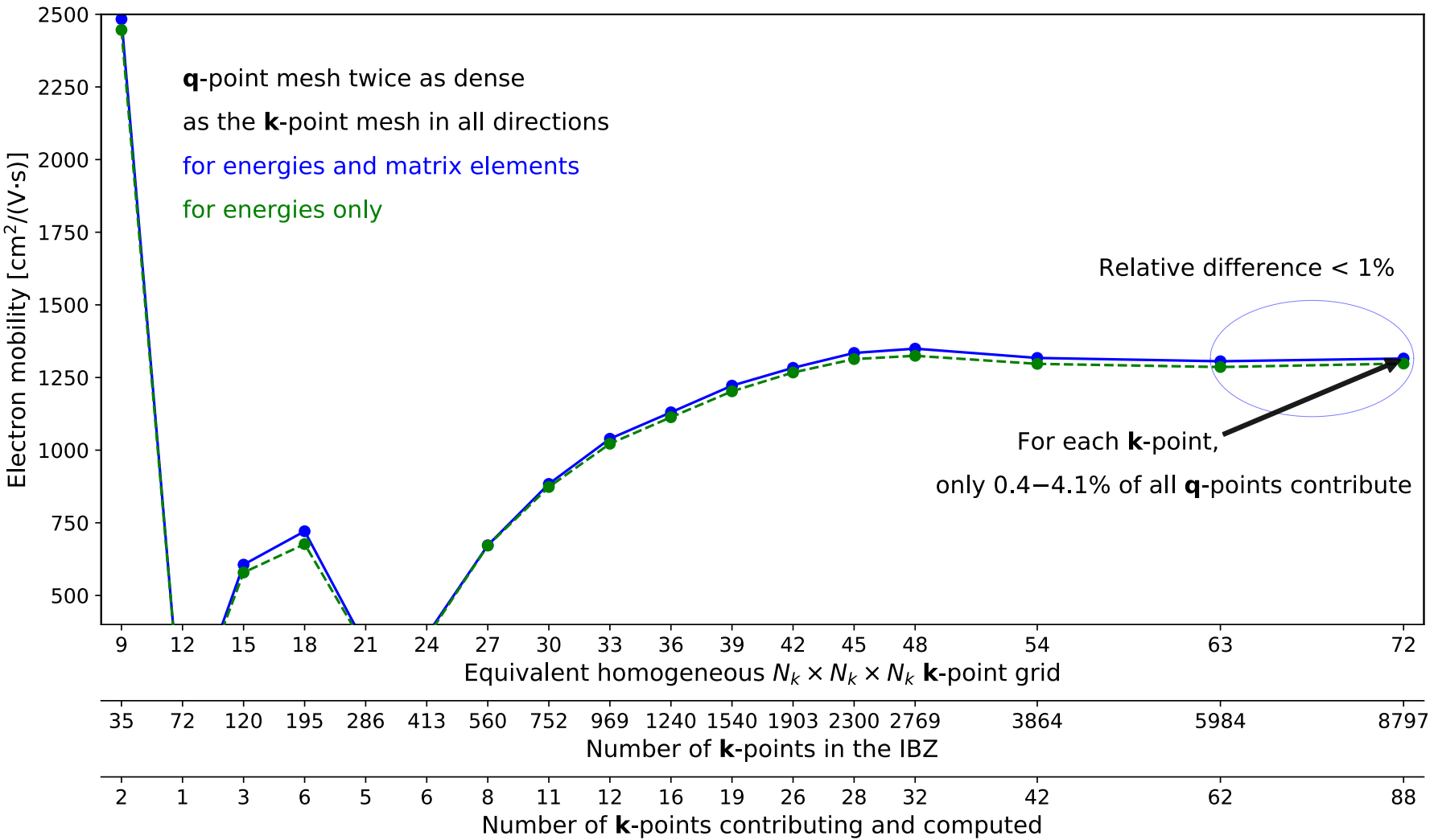
- Selects the important **k**- and **q**-points
- Computes the lifetimes and velocities for these points
- Performs the integration to obtain the phonon-limited mobility



Mobility of electrons in Silicon with ABINIT

New implementation in ABINIT :

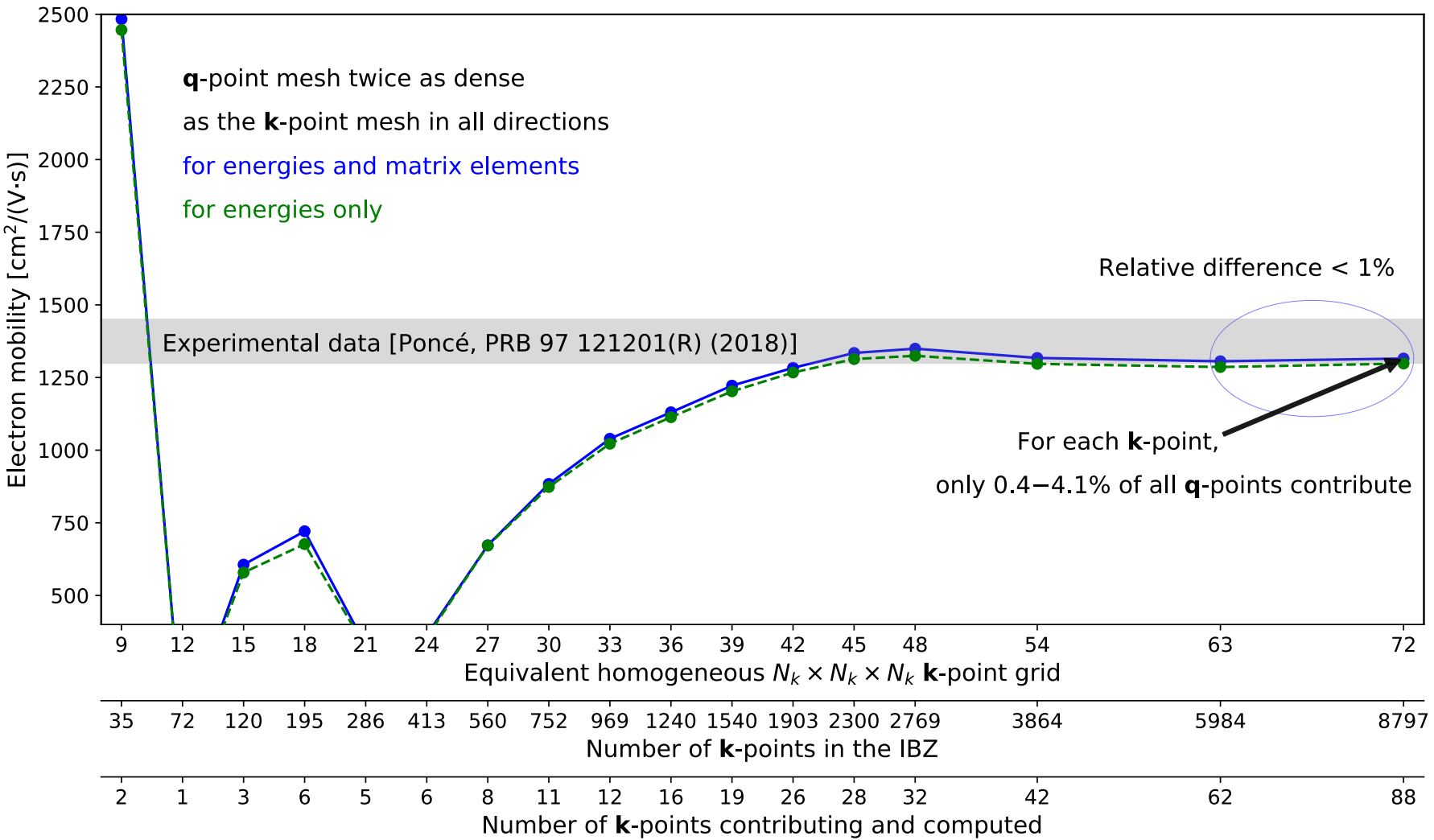
- Selects the important **k**- and **q**-points
- Computes the lifetimes and velocities for these points
- Performs the integration to obtain the phonon-limited mobility



Mobility of electrons in Silicon with ABINIT

New implementation in ABINIT :

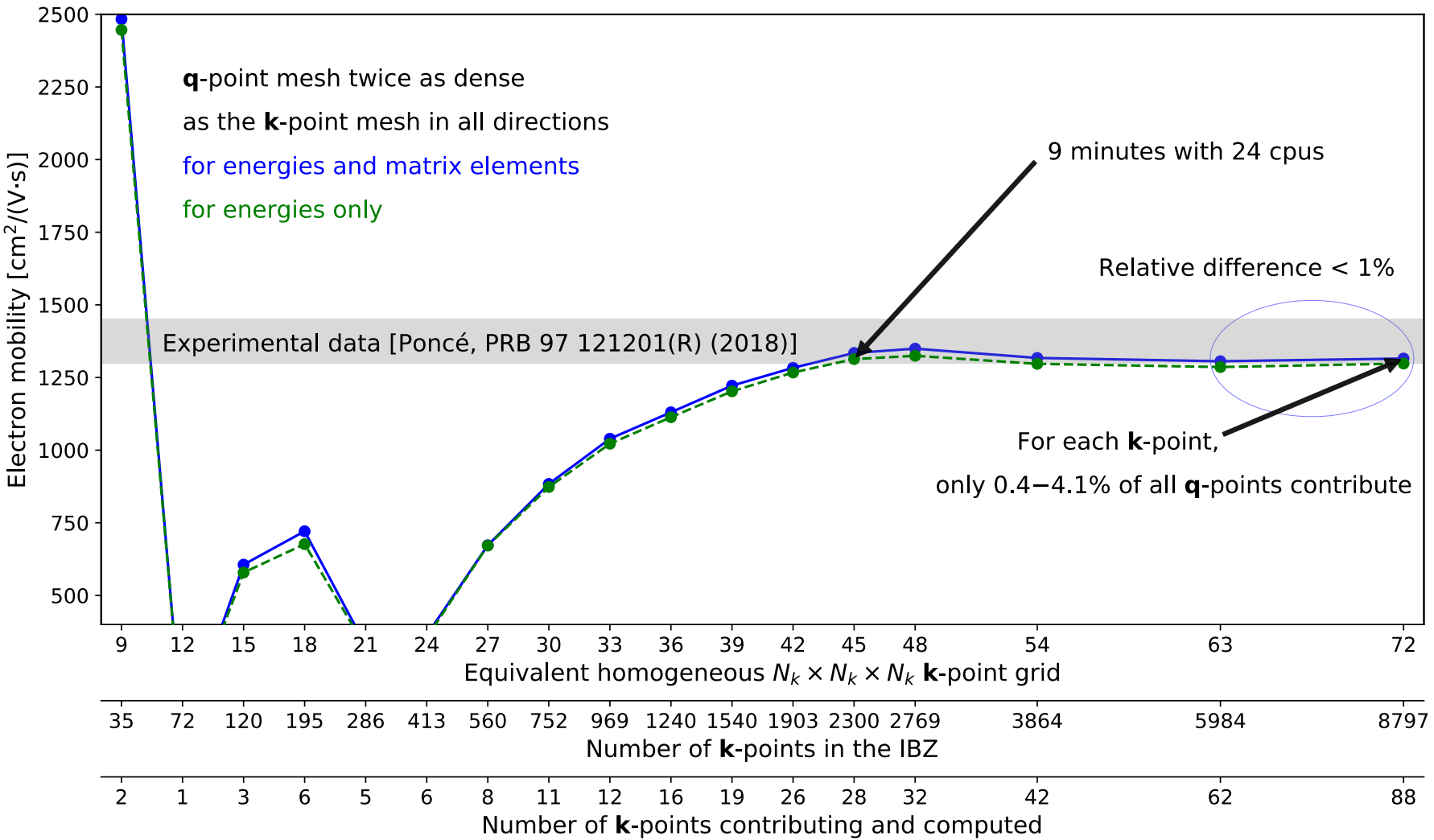
- Selects the important **k**- and **q**-points
- Computes the lifetimes and velocities for these points
- Performs the integration to obtain the phonon-limited mobility



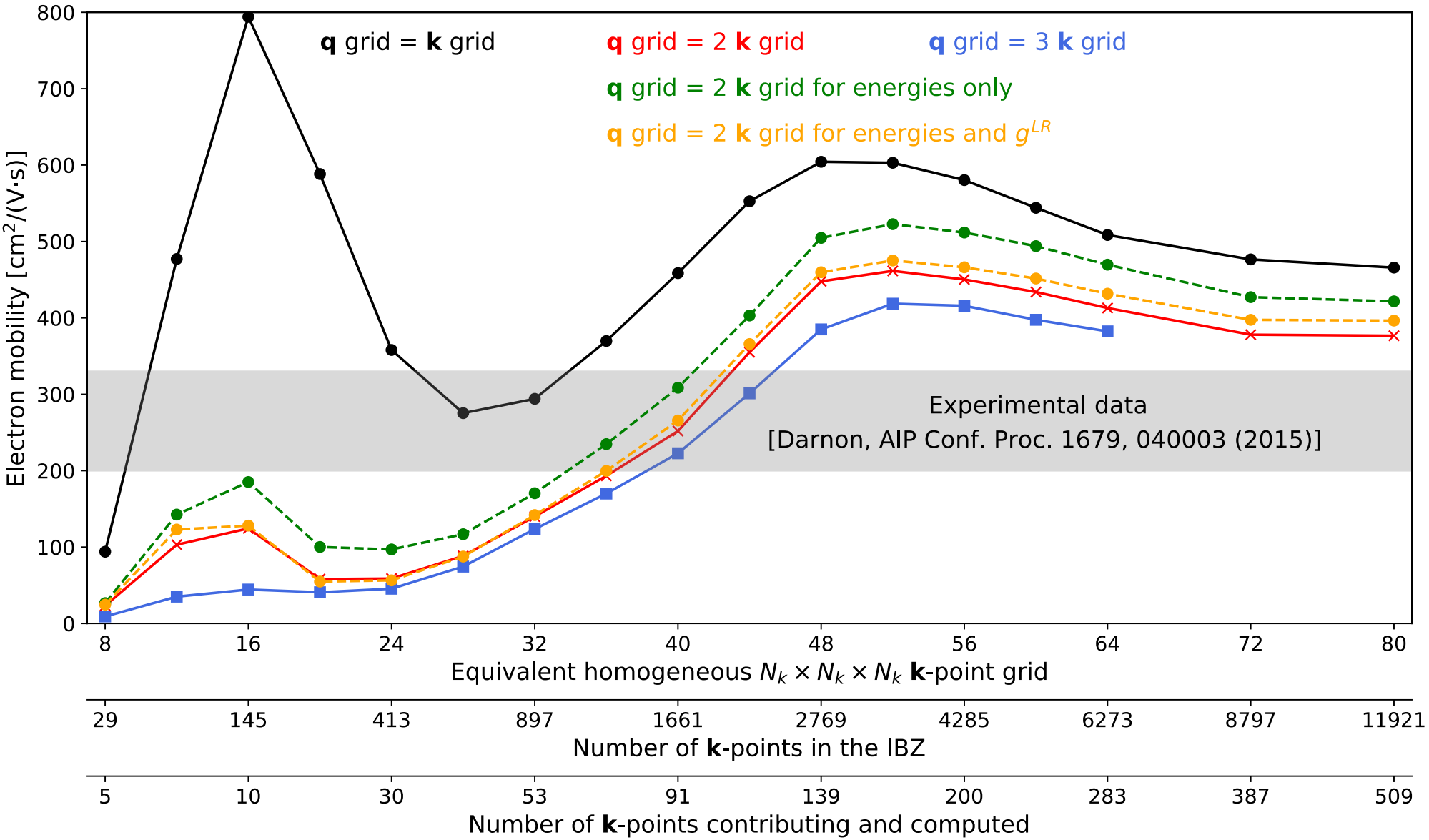
Mobility of electrons in Silicon with ABINIT

New implementation in ABINIT :

- Selects the important **k**- and **q**-points
- Computes the lifetimes and velocities for these points
- Performs the integration to obtain the phonon-limited mobility



Mobility of electrons in GaP with ABINIT



First part : conclusion

- We can compute lifetimes with ABINIT, without using Wannier functions or atomic orbitals
  Note : this was already possible in previous versions of ABINIT
- Use of the Tetrahedron method :
  - correct behavior without having to deal with a broadening parameter
  - faster convergence w.r.t. **q** points
- Double-grid technique :
  - the δ functions require denser grid than the $g_{mn,\nu}(\mathbf{k},\mathbf{q})$
  - the diverging part of the matrix elements can be computed on a denser grid
- Phonon-limited mobility : we compute only what is necessary
  - $\tau_{n\mathbf{k}}$ for few % of **k** points, $g_{mn,\nu}(\mathbf{k},\mathbf{q})$ for few % of **q** points
- Second part : by Henrique Miranda

Electron-phonon self-energy using plane waves

$$\Sigma_{n\mathbf{k}}(\omega, \varepsilon_F, T) = \sum_{m,\nu} \int_{BZ} \frac{d\mathbf{q}}{\Omega_{BZ}} |g_{mn,\nu}(\mathbf{k}, \mathbf{q})|^2 \left[\frac{n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}} + i\gamma} + \frac{n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}} + i\gamma} \right]$$

$$g_{mn,\nu}(\mathbf{k}, \mathbf{q}) = \langle \psi_{m\mathbf{k}+\mathbf{q}} | \Delta_{\nu\mathbf{q}} V^{KS} | \psi_{n\mathbf{k}} \rangle$$

Electron-phonon self-energy using plane waves

$$\Sigma_{n\mathbf{k}}(\omega, \varepsilon_F, T) = \sum_{m,\nu} \int_{BZ} \frac{d\mathbf{q}}{\Omega_{BZ}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 \left[\frac{n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}} + i\gamma} + \frac{n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}} + i\gamma} \right]$$

$$g_{mn,\nu}(\mathbf{k},\mathbf{q}) = \langle \psi_{m\mathbf{k}+\mathbf{q}} | \Delta_{\nu\mathbf{q}} V^{KS} \psi_{n\mathbf{k}} \rangle$$

- Interpolate the DFPT potentials
  - Remove and add the long-range part (Fröhlich)
  - Reduce memory for potentials: boxcutmin, single precision, distribution over perturbations and q-points

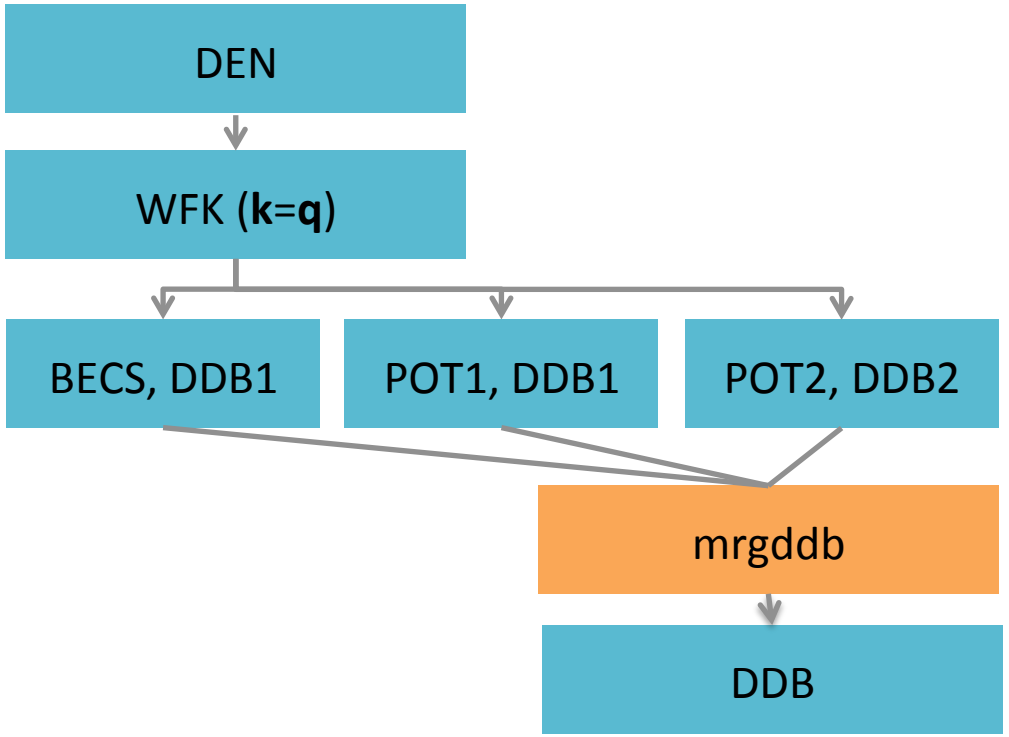
Electron-phonon self-energy using plane waves

$$\Sigma_{n\mathbf{k}}(\omega, \varepsilon_F, T) = \sum_{m,\nu} \int_{BZ} \frac{d\mathbf{q}}{\Omega_{BZ}} |g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2 \left[\frac{n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}} + i\gamma} + \frac{n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}} + i\gamma} \right]$$

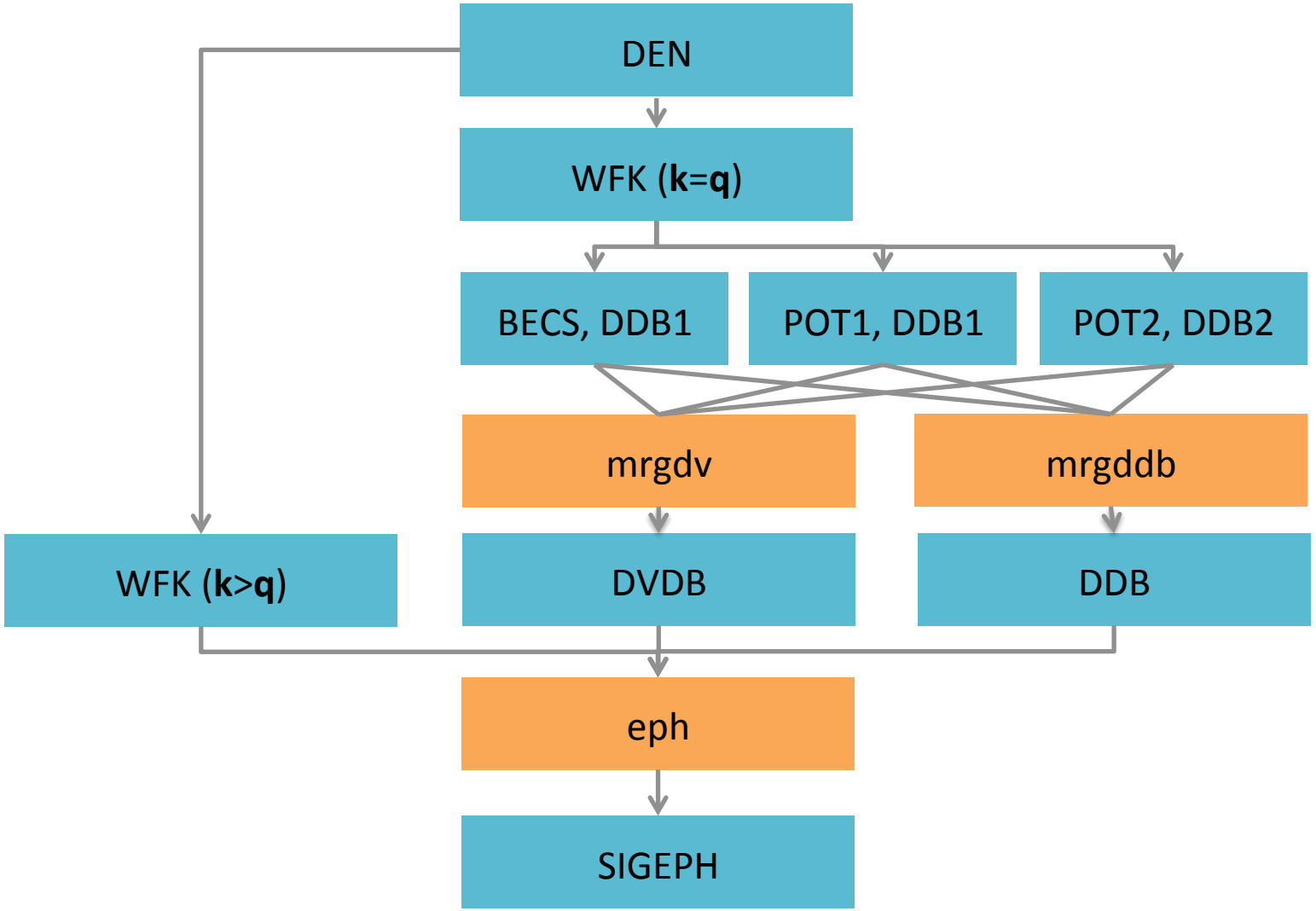
$$g_{mn,\nu}(\mathbf{k},\mathbf{q}) = \langle \psi_{m\mathbf{k}+\mathbf{q}} | \Delta_{\nu\mathbf{q}} V^{KS} | \psi_{n\mathbf{k}} \rangle$$

- Interpolate the DFPT potentials
  - Remove and add the long-range part (Fröhlich)
  - Reduce memory for potentials: boxcutmin, single precision, distribution over perturbations and q-points
- Wave functions from NSCF calculation
- Compute $g_{mn,\nu}(\mathbf{k},\mathbf{q})$ and accumulate on-the-fly (avoid IO)

Electron-phonon self-energy workflow



Electron-phonon self-energy workflow



Checkpoint and Restart

- Long calculations -> restart feature!
- Additional array with “done”/”not done” written inside the loop over k-points
- If the calculation is crashed, restart with

```
eph_restart 1
```

- Save computational time!

Bolzmann transport in the SERTA

- Mobility in the linearized Boltzmann equation:

$$\mu_{\alpha\beta}(\varepsilon_F, T) = \frac{e}{n(\varepsilon_F, T)\Omega} \sum_n \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \left(-\frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon} \right)$$

Bolzmann transport in the SERTA

- Mobility in the linearized Boltzmann equation:

$$\mu_{\alpha\beta}(\varepsilon_F,T)=\frac{e}{n(\varepsilon_F,T)\Omega}\sum_n\int\frac{d\mathbf{k}}{\Omega_{BZ}}\mathbf{v}_{n\mathbf{k},\alpha}\mathbf{v}_{n\mathbf{k},\beta}\tau_{n\mathbf{k}}(\varepsilon_F,T)\left(-\frac{\partial f(\varepsilon_{n\mathbf{k}},\varepsilon_F,T)}{\partial\varepsilon}\right)$$

- Lifetimes from imaginary part of self-energy:

$$\Sigma_{n\mathbf{k}}(\omega,\varepsilon_F,T)=\sum_{m,\nu}\int_{BZ}\frac{d\mathbf{q}}{\Omega_{BZ}}|g_{mn,\nu}(\mathbf{k},\mathbf{q})|^2\left[\frac{n_{\nu\mathbf{q}}+f_{m\mathbf{k}+\mathbf{q}}}{\omega-\varepsilon_{m\mathbf{k}+\mathbf{q}}+\omega_{\nu\mathbf{q}}+i\gamma}+\frac{n_{\nu\mathbf{q}}+1-f_{m\mathbf{k}+\mathbf{q}}}{\omega-\varepsilon_{m\mathbf{k}+\mathbf{q}}-\omega_{\nu\mathbf{q}}+i\gamma}\right]$$

$$\frac{1}{\tau_{n\mathbf{k}}}=\pi\sum_{m,\nu}\int_{\mathrm{BZ}}\frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}}|g_{mn\mathbf{k},\nu\mathbf{q}}|^2\times[(n_{\nu\mathbf{q}}+f_{m\mathbf{k}+\mathbf{q}})\delta(\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}+\omega_{\nu\mathbf{q}})+(n_{\nu\mathbf{q}}+1-f_{m\mathbf{k}+\mathbf{q}})\delta(\varepsilon_{n\mathbf{k}}-\varepsilon_{m\mathbf{k}+\mathbf{q}}-\omega_{\nu\mathbf{q}})]$$

Speeding up imaginary part of SE

Double grid

Eigenvalues dense grid:

1. SKW interpolation
`bs_interp_kmult`
2. NSCF calculation
`tolwfr 1e-15`
`irdwfkfine`
`getwfkfine`

Speeding up imaginary part of SE

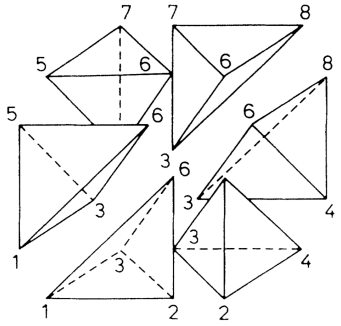
Double grid

Eigenvalues dense grid:

1. SKW interpolation
`bs_interp_kmult`
2. NSCF calculation
`tolwfr 1e-15`
`irdwfkfine`
`getwfkfine`

Tetrahedron integration

`eph_intmeth 2`



Speeding up imaginary part of SE

Double grid

Eigenvalues dense grid:

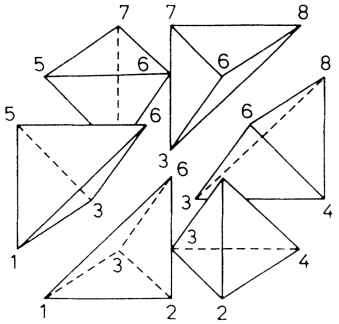
1. SKW interpolation
   `bs_interp_kmult`
2. NSCF calculation
   `tolwfr 1e-15`

`irdwfkfine`

`getwfkfine`

Tetrahedron integration

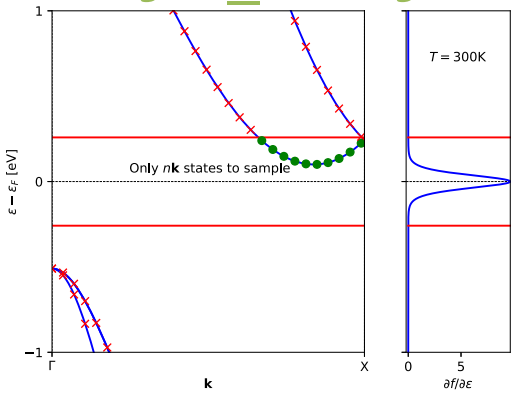
`eph_intmeth 2`



Q and K-point filtering

`kerange`

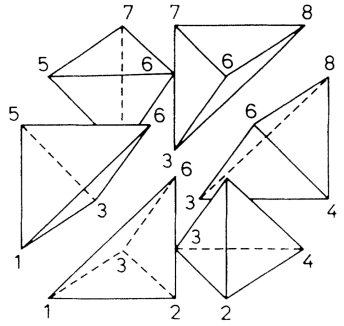
`sigma_erange`



Speeding up imaginary part of SE

Tetrahedron integration

`eph_intmeth 2`



Tetrahedron method

$$\Sigma_{n\mathbf{k}}(\omega, \varepsilon_F, T) = \sum_{m,\nu} \int_{BZ} \frac{d\mathbf{q}}{\Omega_{BZ}} |g_{mn,\nu}(\mathbf{k}, \mathbf{q})|^2 \left[\frac{n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}} + i\gamma} + \frac{n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}}}{\omega - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}} + i\gamma} \right]$$

zcut small to reproduce the limit,

large enough to avoid numeric problems

Double convergence γ and q

Tetrahedron method

$$\frac{1}{\tau_{n\mathbf{k}}} = \pi \sum_{m,\nu} \int_{\mathrm{BZ}} \frac{d\mathbf{q}}{\Omega_{\mathrm{BZ}}} |g_{mn\mathbf{k},\nu\mathbf{q}}|^2 \times [(n_{\nu\mathbf{q}} + f_{m\mathbf{k}+\mathbf{q}})\delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} + \omega_{\nu\mathbf{q}}) +$$

$$(n_{\nu\mathbf{q}} + 1 - f_{m\mathbf{k}+\mathbf{q}})\delta(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}+\mathbf{q}} - \omega_{\nu\mathbf{q}})]$$

zcut small to reproduce the limit,

large enough to avoid numeric problems

Double convergence γ and q

tetrahedron linear interpolation of the eigenvalues and matrix elements

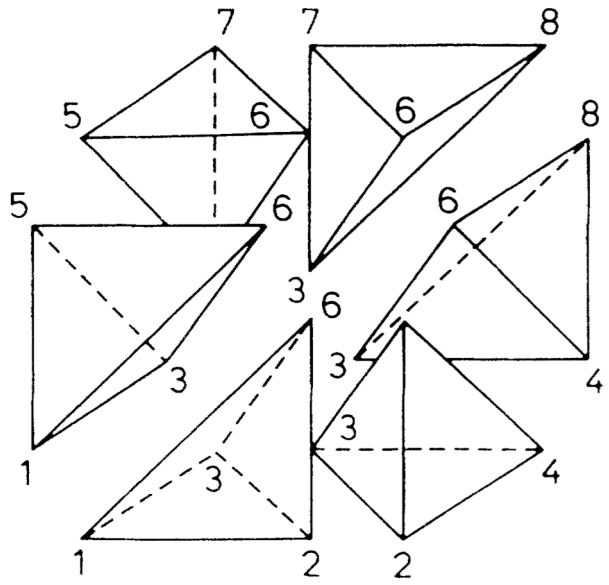
Converge q only

Tetrahedron method 101

- Tessellate the Brillouin Zone using tetrahedrons

$$\pi I(\omega) = \int_{BZ} F_{\mathbf{k}} \delta(\omega - \varepsilon_{\mathbf{k}})$$

$$\pi I(\omega) = \Omega_{\text{tetra}} \sum_{i=1}^{N_{\text{tetra}}} g_i(\omega) \sum_{s=1}^{4} I_s^i(\omega) F_s^i$$



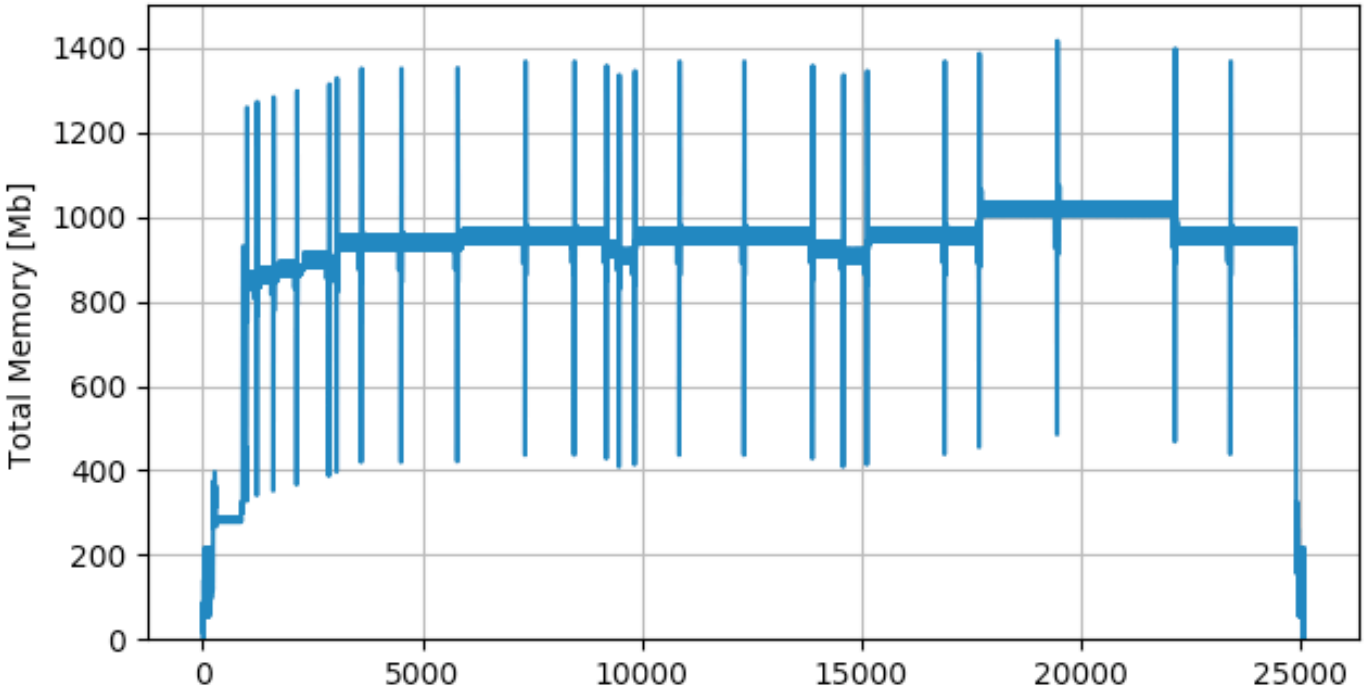
- Simple analytical expressions depend on the energies at the summits

[1] A.H. MacDonald, S.H. Vosko, and P.T. Coleridge, Journal of Physics C: Solid State Physics **12**, 2991 (1979)
[2] P.E. Blöchl, O. Jepsen, and O.K. Andersen, Phys. Rev. B **49**, 16223 (1994)
Many others... Also dicussion with Atsushi Togo

Old tetrahedron implementation

1. Create list of all tetrahedra
2. Hash and sort
3. Reduce to unique

(Double memory allocation)



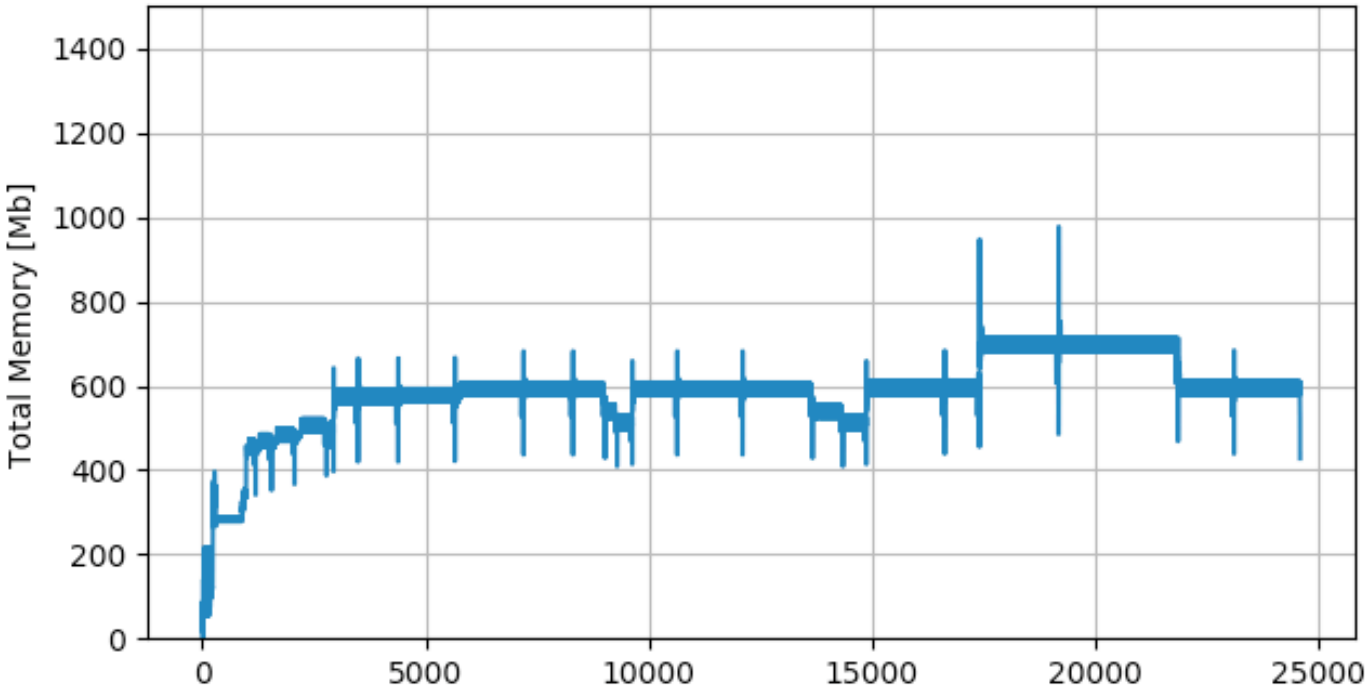
memory trace
abinit --abimem-level 3
abimem.py plot <file>

GaAs
72x72x72
k and q grid

New tetrahedron implementation

1. Generate 1 tetrahedron and hash
2. If new store, if already exists add multiplicity
3. Repeat for all tetrahedra

(Lower memory footprint)



memory trace
abinit --abimem-level 3
abimem.py plot <file>

GaAs
72x72x72
k and q grid

Bolzmann transport in the SERTA

- Mobility in the linearized Boltzmann equation:

$$\mu_{\alpha\beta}(\varepsilon_F, T) = \frac{e}{n(\varepsilon_F, T)\Omega} \sum_n \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \left(-\frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon} \right)$$

Group velocity matrix elements

$$\mathbf{v}_{n\mathbf{k},\alpha} = \langle \psi_{n\mathbf{k}} | \partial_{\mathbf{k},\alpha} H | \psi_{n\mathbf{k}} \rangle$$

- Before: DFPT run on FBZ for each 3 directions and read diagonal matrix elements from 1WF files
  - Lots of IO and waste!

Group velocity matrix elements

$$\mathbf{v}_{n\mathbf{k},\alpha} = \langle \psi_{n\mathbf{k}} | \hat{p}_\alpha + i[V_{NL}, \hat{r}_\alpha] | \psi_{n\mathbf{k}} \rangle$$

- Before: DFPT run on FBZ for each 3 directions and read diagonal matrix elements from 1WF files
  - Lots of IO and waste!
- Compute on-the-fly using `nc_ihr_comm` (chi in GW)
  - Only works for norm-conserving
  - Not most efficient for off-diagonal matrix elements

Group velocity matrix elements

$$\mathbf{v}_{n\mathbf{k},\alpha} = \langle \psi_{n\mathbf{k}} | \partial_{\mathbf{k},\alpha} H | \psi_{n\mathbf{k}} \rangle$$

DFPT routines:

- load_spin_hamiltonian
- load_spin_rf_hamiltonian
- getgh1c_setup
- getgh1c

Group velocity matrix elements

$$\mathbf{v}_{n\mathbf{k},\alpha} = \langle \psi_{n\mathbf{k}} | \partial_{\mathbf{k},\alpha} H | \psi_{n\mathbf{k}} \rangle$$

DFPT routines:

- load_spin_hamiltonian
- load_spin_rf_hamiltonian
- getgh1c_setup
- getgh1c

Object dkkop_t:

- ddkop_setup_spin_kpoint
- ddkop_apply
- ddkop_get_velocity

Can be reused in different contexts: transport, chi, interpolation, etc…

WFs in memory -> recomputing better than IO

Same for EPH matrix elements

Transport computation driver

$$\mu_{\alpha\beta}(\varepsilon_F, T) = \frac{e}{n(\varepsilon_F, T)\Omega} \sum_n \int \frac{d\mathbf{k}}{\Omega_{BZ}} \mathbf{v}_{n\mathbf{k},\alpha} \mathbf{v}_{n\mathbf{k},\beta} \tau_{n\mathbf{k}}(\varepsilon_F, T) \left(-\frac{\partial f(\varepsilon_{n\mathbf{k}}, \varepsilon_F, T)}{\partial \varepsilon} \right)$$

Transport computation driver

$$\mu_{\alpha\beta}(\varepsilon_F,T)=\frac{e}{n(\varepsilon_F,T)\Omega}\sum_n\int\frac{d\mathbf{k}}{\Omega_{BZ}}\mathbf{v}_{n\mathbf{k},\alpha}\mathbf{v}_{n\mathbf{k},\beta}\tau_{n\mathbf{k}}(\varepsilon_F,T)\left(-\frac{\partial f(\varepsilon_{n\mathbf{k}},\varepsilon_F,T)}{\partial\varepsilon}\right)$$

$$\mu_{\alpha\beta}(\varepsilon_F,T)=\frac{e}{n(\varepsilon_F,T)\Omega}\int K(\omega,\varepsilon_F,T)\left(-\frac{\partial f(\omega,\varepsilon_F,T)}{\partial\varepsilon}\right)d\omega$$

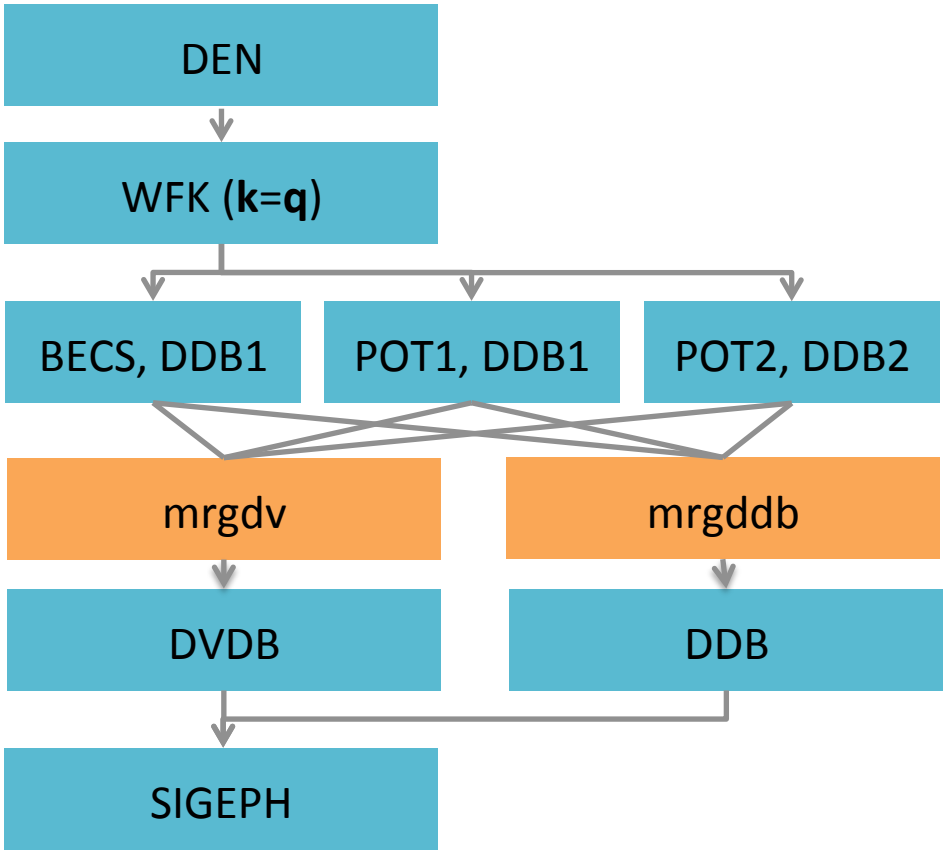
$$K(\omega,\varepsilon_F,T)=\sum_n\int\frac{d\mathbf{k}}{\Omega_{BZ}}\mathbf{v}_{n\mathbf{k},\alpha}\mathbf{v}_{n\mathbf{k},\beta}\tau_{n\mathbf{k}}(\varepsilon_F,T)\delta(\omega-\varepsilon_{n\mathbf{k}})$$

- Can use the tetrahedron method!

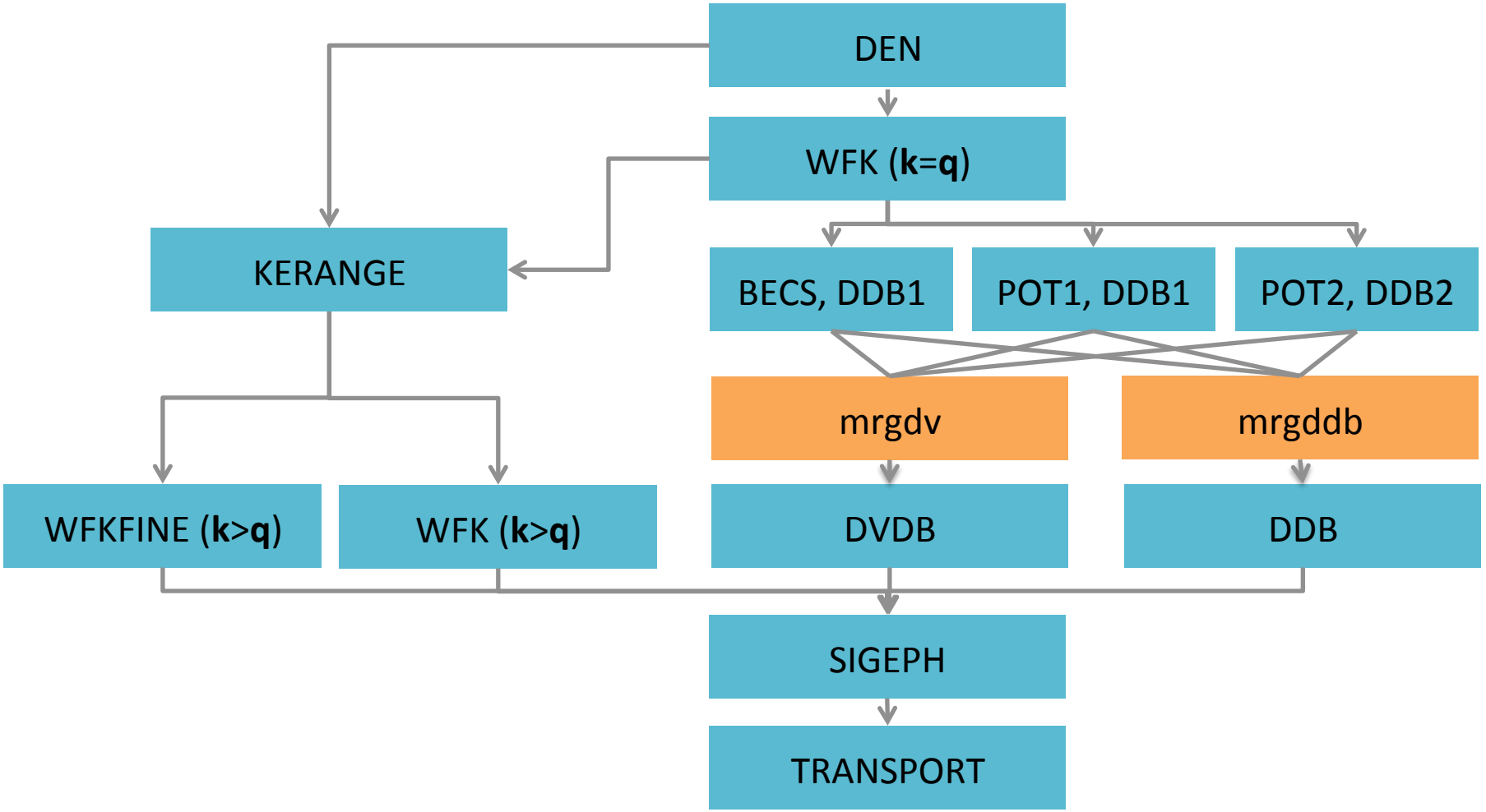
Transport computation driver

- New transport computation driver starting from `*_SIGEPH.nc` file
- Computed automatically after sigmaph when `eph_task -4`
- Compute conductivity, mobility, Seebeck.
- Write to `*_TRANSPORT.nc` netcdf file
- Analyze results using Abipy

Transport computation workflow



Transport computation workflow



Unit tests

- Test individual routines outside of their normal scope `95_drive/m_unittests.F90`
- Two use cases so far:
  - New tetrahedron routines (real part of SE)
  - Implement `symkpt` and `listkk` routines with better scaling for our applications
- Work in progress…

Lessons learned

- **IO is bad!** Lots of IO will quickly become the bottleneck of the computation
- **FLOPS are cheap** Re-computing is better than IO

Lessons learned

- **IO is bad!** Lots of IO will quickly become the bottleneck of the computation
- **FLOPS are cheap** Re-computing is better than IO
- **Treat integrals nicely** When performing difficult BZ integrals prefer tetrahedron (or better)

Lessons learned

- **IO is bad!** Lots of IO will quickly become the bottleneck of the computation
- **FLOPS are cheap** Re-computing is better than IO
- **Treat integrals nicely** When performing difficult BZ integrals prefer tetrahedron (or better)
- **Save memory** Aim to 2GB per core, distribute memory, recompute, avoid unnecessary allocations

Lessons learned

- **IO is bad!** Lots of IO will quickly become the bottleneck of the computation
- **FLOPS are cheap** Re-computing is better than IO
- **Treat integrals nicely** When performing difficult BZ integrals prefer tetrahedron (or better)
- **Save memory** Aim to 2GB per core, distribute memory, recompute, avoid unnecessary allocations
- **Checkpoint is good** Save work, allow restart

Further work

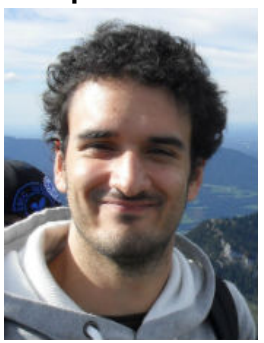
1. Special treatment of Fröhlich matrix elements
2. Improve convergence of mobility with k-points using quadratic tetrahedron method
3. Faster and scalable k-points machinery: `symkpt` and `listkk`

Acknowledgments

Geoffroy Hautier

Gian-Marco Rignanese

Henrique Miranda



Matteo Giantomassi

Funding :

FREEDOM TO RESEARCH

UCLouvain
Institute of Condensed Matter and Nanosciences

Thank you!