First experiences with overlap fermions based on the Brillouin kernel

Stephan Dürr

University of Wuppertal
Jülich Supercomputing Center

work with Giannis Koutsou (Castorc, Cyprus)

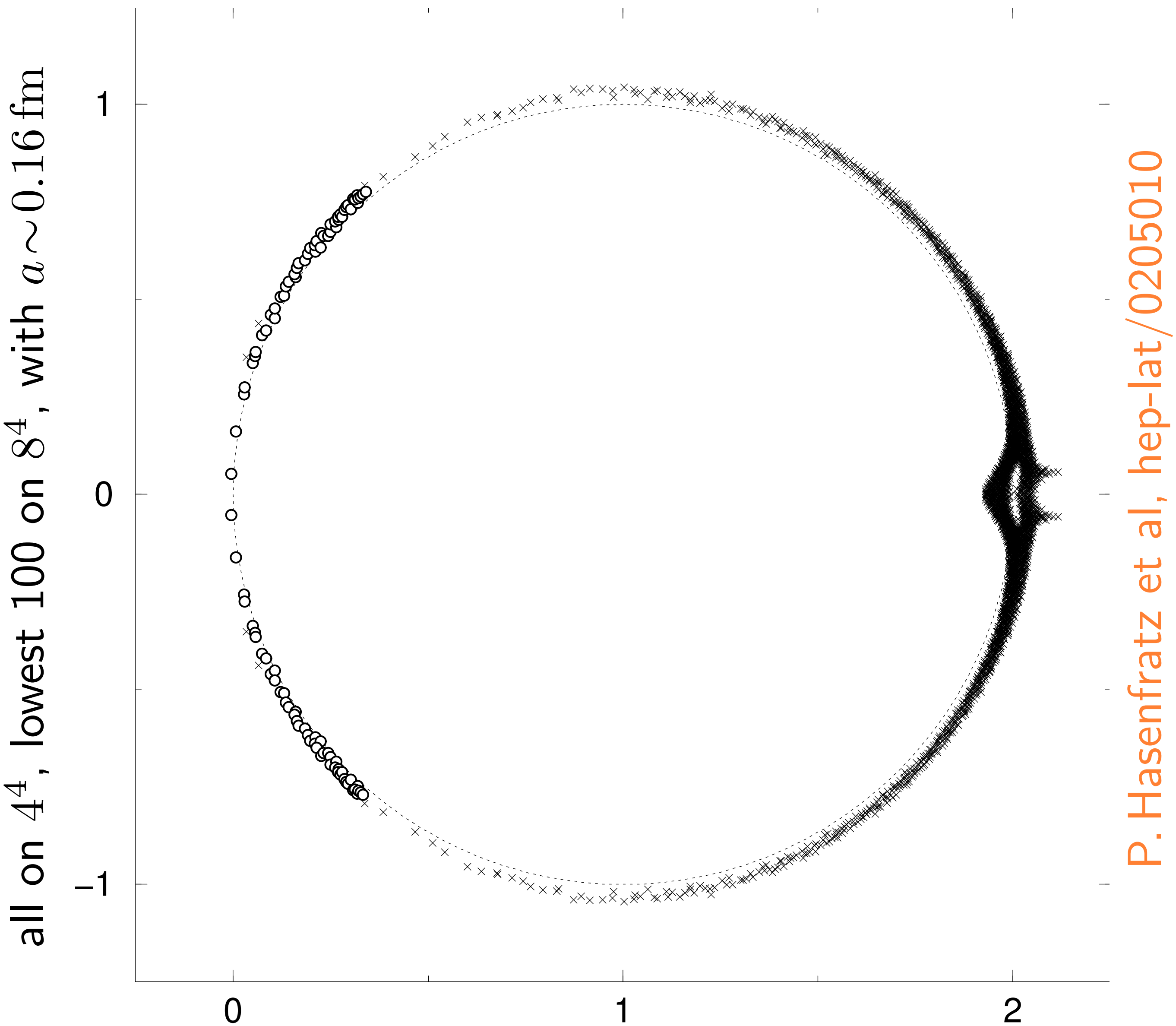
Lattice 2016 – Chiral Symmetry – 25. July 2016

In memoriam Peter Hasenfratz and Keisuke Jimmy Juge

Peter Hasenfratz (22.9.1946 – 9.4.2016)

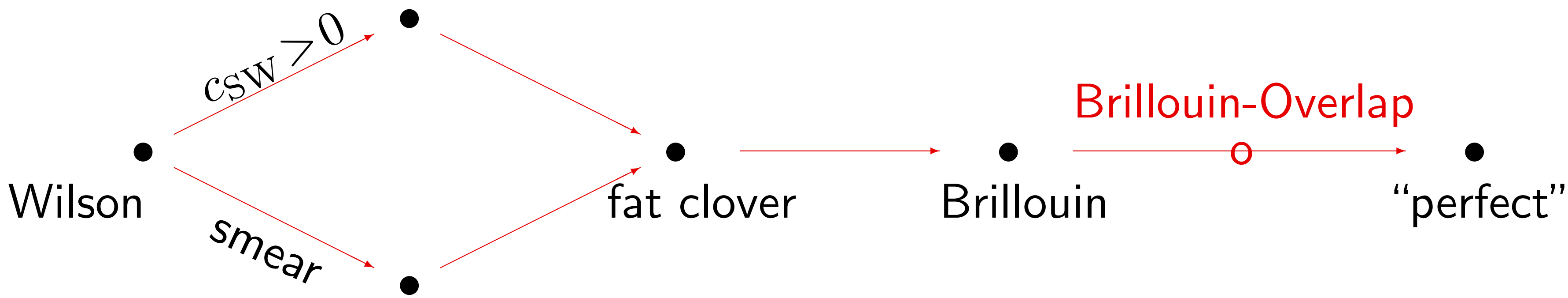
Keisuke Jimmy Juge (21.9.1971 – 30.6.2016)

The legacy of the "perfect action"



Perfect action: Ginsparg-Wilson spectrum, linear dispersion relation, no cut-off effects !

- **Pedestrian (bottom-up) approach**



$$D(x,y)=\frac{1}{2}\sum_\mu\Big\{(\gamma_\mu-I)U_\mu(x)\delta_{x+\hat\mu,y}-(\gamma_\mu+I)U_\mu^\dagger(x-\hat\mu)\delta_{x-\hat\mu,y}\Big\}+\frac{1}{2\kappa}\delta_{x,y}-\frac{c_{SW}}{2}...$$

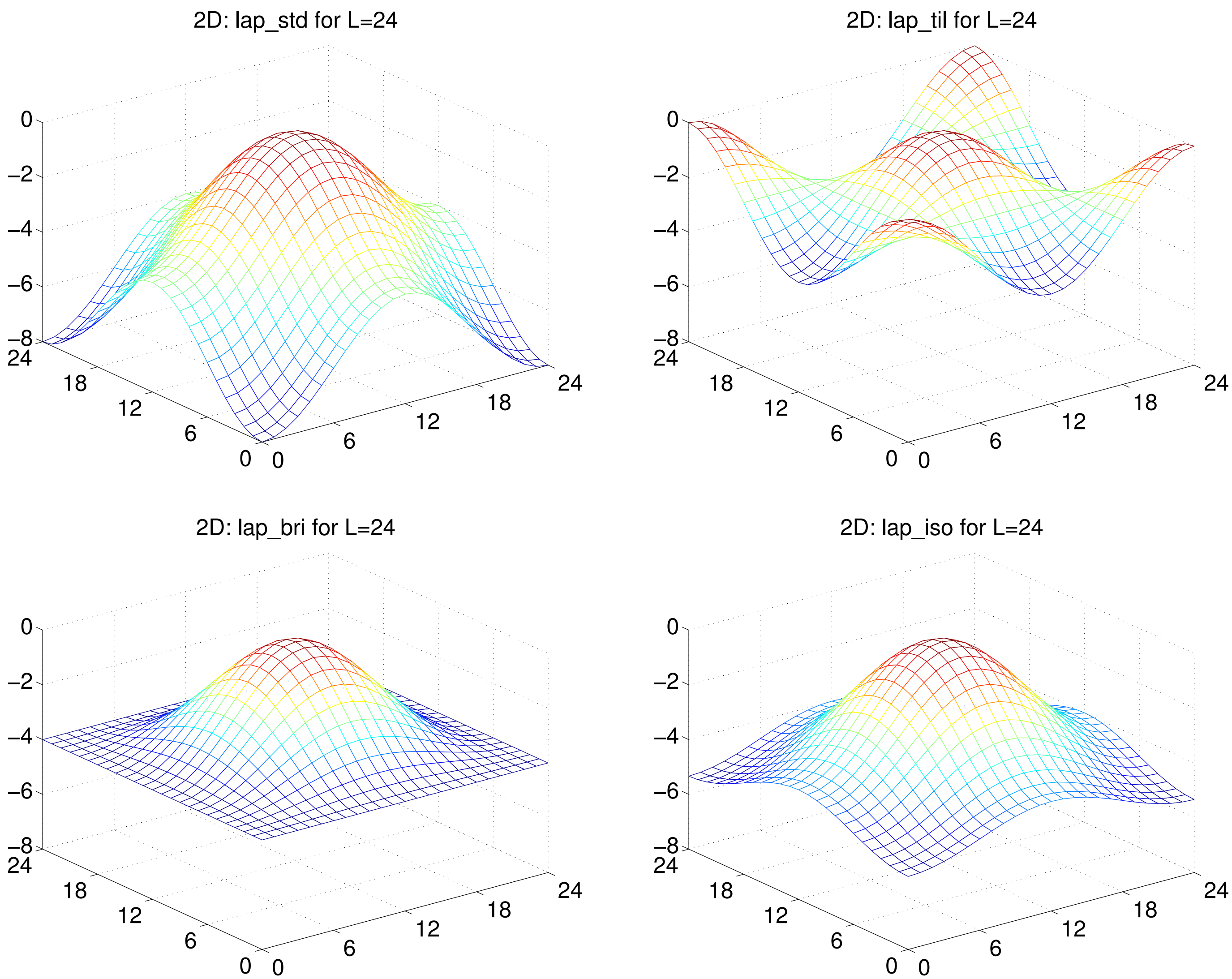
with $\sigma_{\mu\nu}=\frac{i}{2}[\gamma_\mu,\gamma_\nu]$ and $F_{\mu\nu}$ the hermitean clover-leaf field-strength tensor equals

$$D(x,y)=\sum_\mu\gamma_\mu\nabla_\mu^{std}(x,y)-\frac{a}{2}\triangle^{std}(x,y)+m_0\delta_{x,y}+\text{improvement term}$$

with the two mass parameters relating to each other through $1/(2\kappa)=4+am_0$.

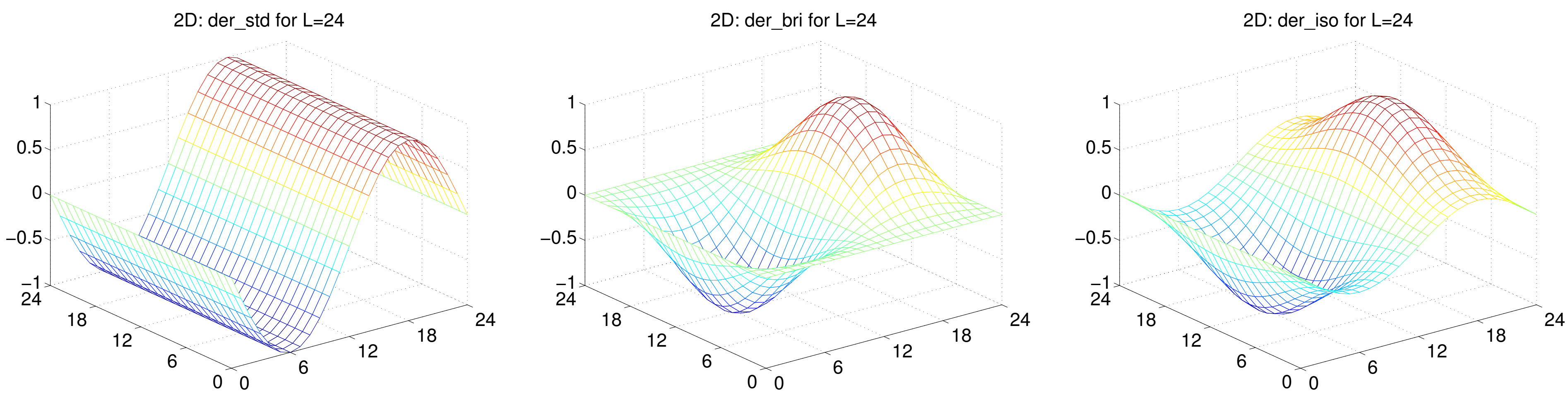
Brillouin operator in 2D

- **4 options for Laplacian**



Standard Laplacian in 2D: $\hat{\triangle} = 2\cos(k_1) + 2\cos(k_2) - 4$
Tilted Laplacian in 2D: $\hat{\triangle} = 2\cos(k_1)\cos(k_2) - 2$
Brillouin Laplacian in 2D: $\hat{\triangle} = 4\cos^2(k_1/2)\cos^2(k_2/2) - 4$
Isotropic Laplacian in 2D: $\hat{\triangle} = [2\cos(k_1)\cos(k_2) + 4\cos(k_1) + 4\cos(k_2) - 10]/3$

- **3 options for derivative**

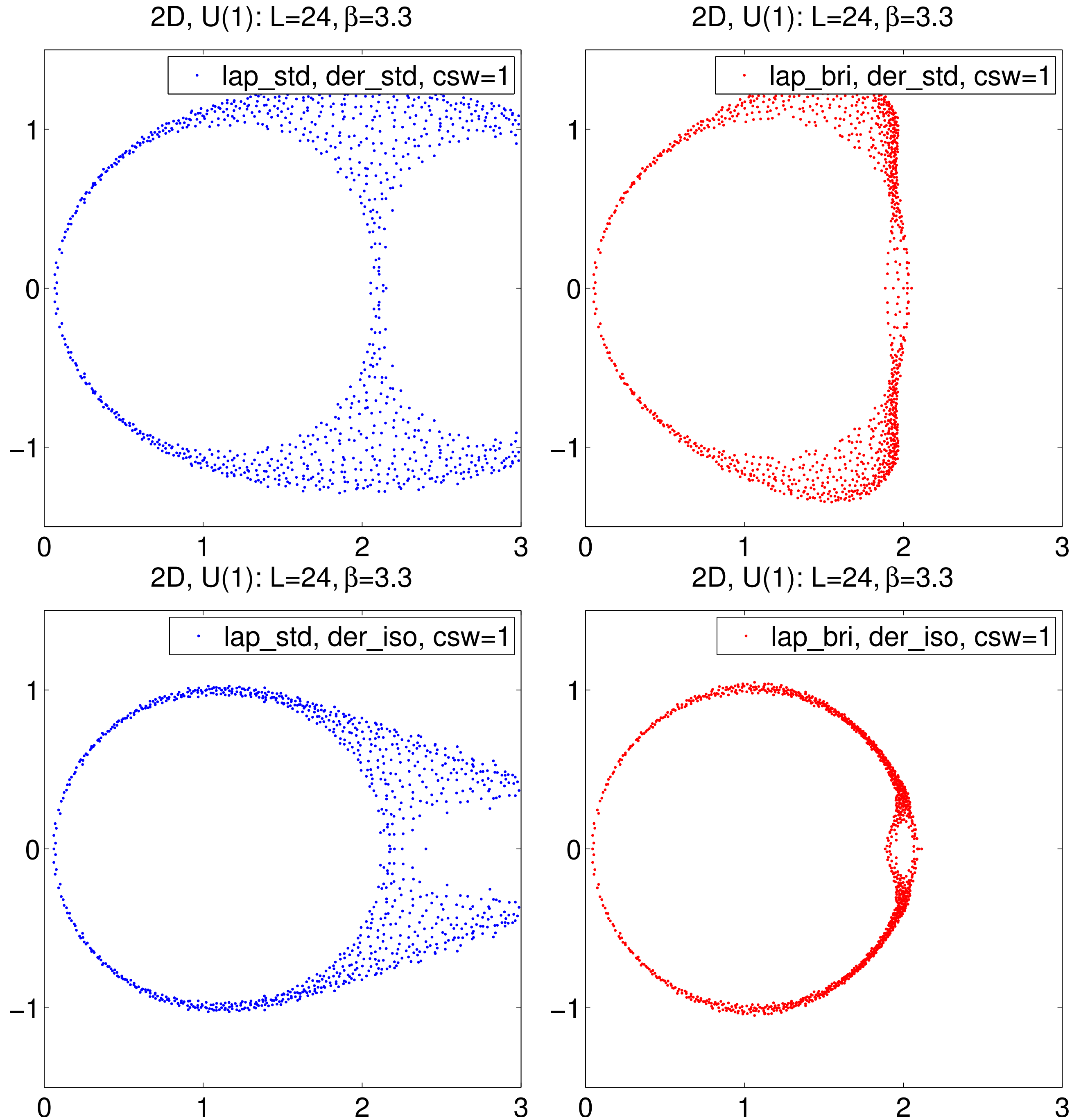


Standard Derivative in 2D: $\hat{\partial}_x = \mathrm{i}\sin(k_1)$
Brillouin Derivative in 2D: $\hat{\partial}_x = \mathrm{i}\sin(k_1)[\cos(k_2) + 1]/2$
Isotropic Derivative in 2D: $\hat{\partial}_x = \mathrm{i}\sin(k_1)[\cos(k_2) + 2]/3$

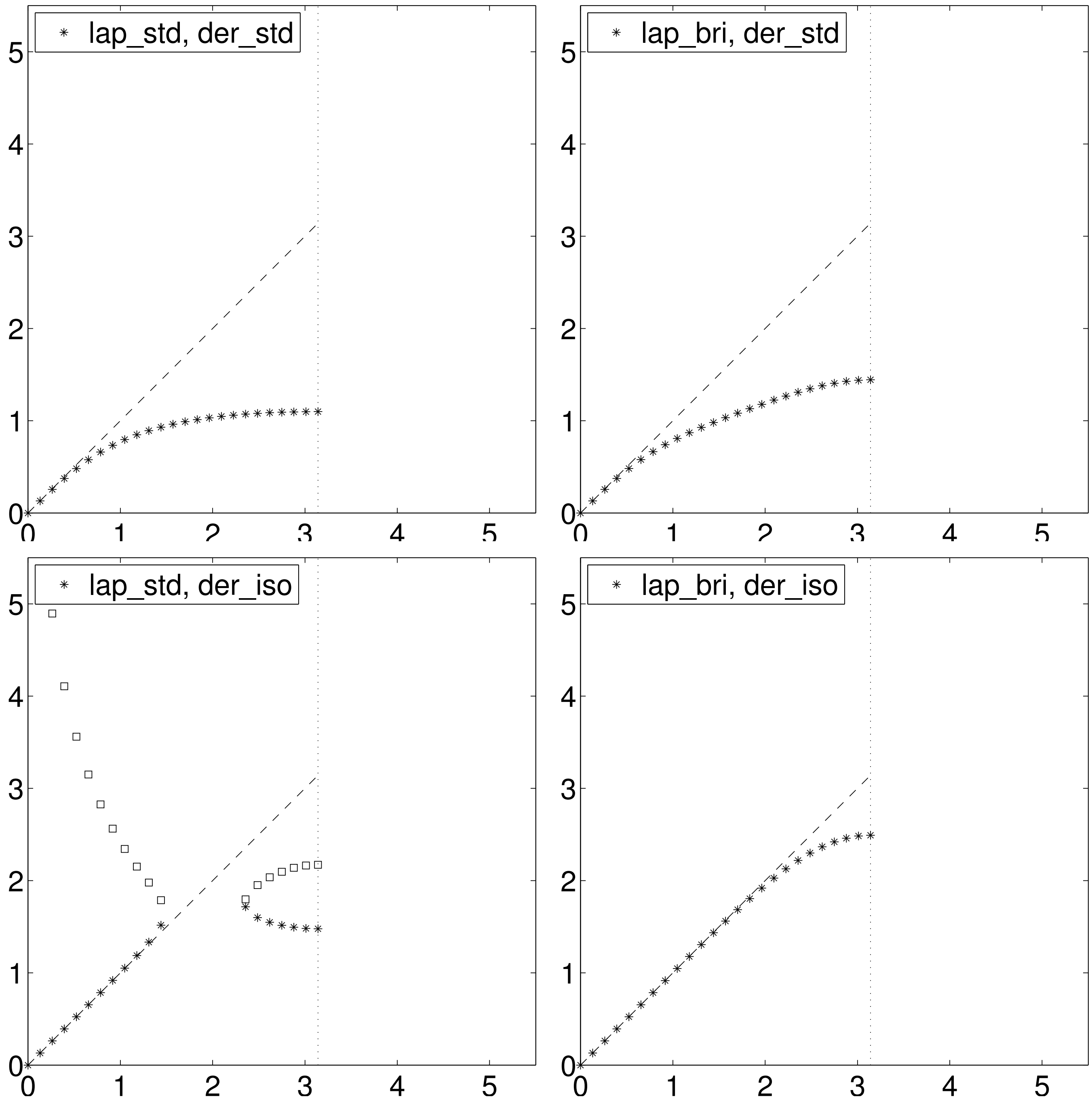
- **Ultralocality**

All 2D stencils (12 options) restricted to $[-1:1]^2$ hypercube (9 sites).

• Eigenvalues of Wilson and Brillouin operators in 2D



• **Free dispersion relation for Wilson and Brillouin operators in 2D**



Brillouin operator in 4D

- **4 options for Laplacian**

Standard Laplacian: $\hat{\triangle} = 2\cos(k_1) + 2\cos(k_2) + 2\cos(k_3) + 2\cos(k_4) - 8$
Tilted Laplacian: $\hat{\triangle} = 2\cos(k_1)\cos(k_2)\cos(k_3)\cos(k_4) - 2$
Brillouin Laplacian: $\hat{\triangle} = 4\cos^2(k_1/2)\cos^2(k_2/2)\cos^2(k_3/2)\cos^2(k_4/2) - 4$
Isotropic Laplacian: $\hat{\triangle} = [2c_1c_2c_3c_4 + 7c_1c_2c_3 + \ldots + 20c_1c_2 + \ldots + 25c_1 + \ldots - 250]/54$

- **3 options for derivative**

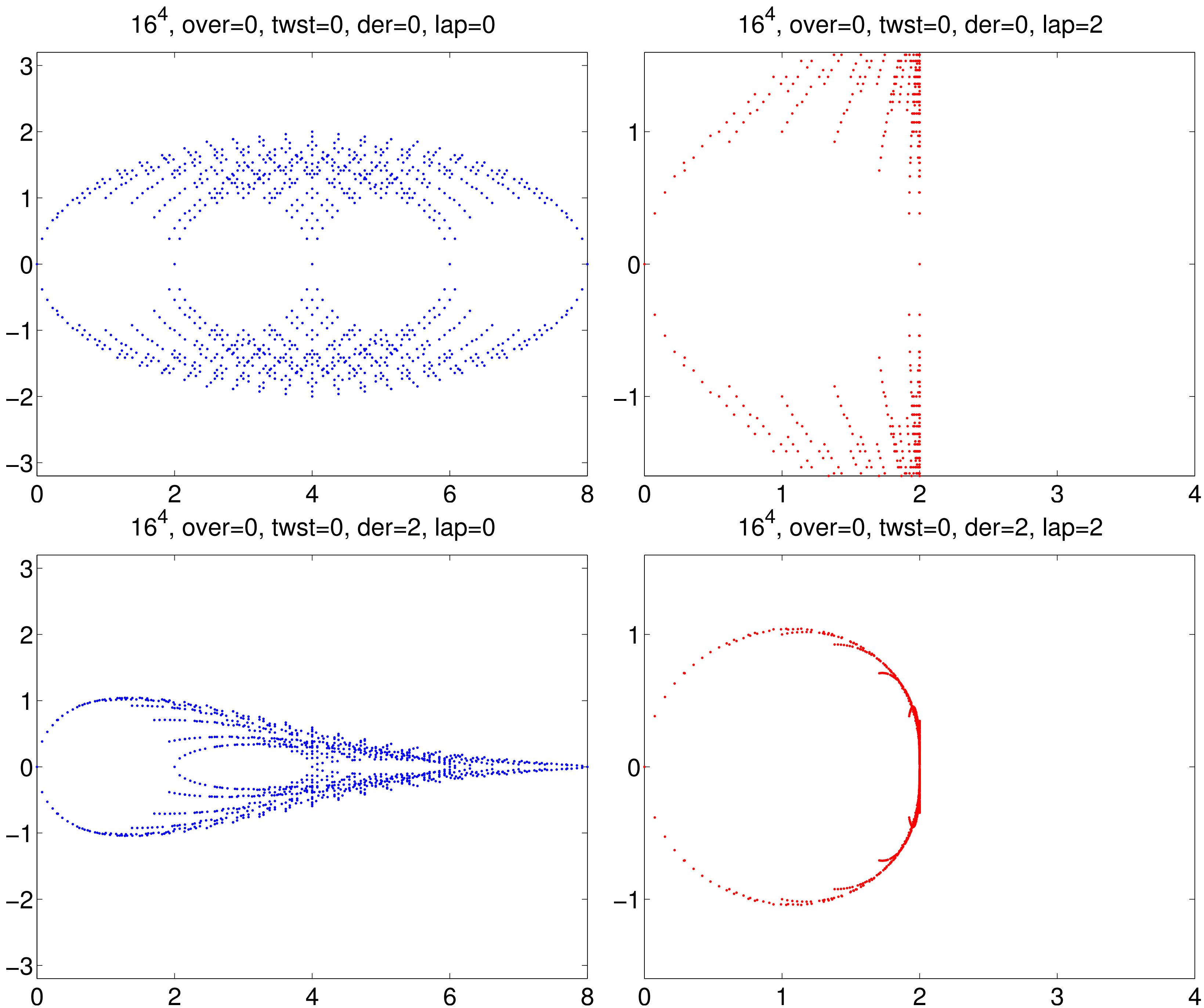
Standard Derivative: $\hat{\partial}_x = \mathrm{i}\sin(k_1)$
Brillouin Derivative: $\hat{\partial}_x = \mathrm{i}\sin(k_1)[\cos(k_2)+1][\cos(k_3)+1][\cos(k_4)+1]/8$
Isotropic Derivative: $\hat{\partial}_x = \mathrm{i}\sin(k_1)[\cos(k_2)+2][\cos(k_3)+2][\cos(k_4)+2]/27$

(*4D stencils look similar to 2D case*)

- **Ultralocality**

All 4D stencils (12 options) restricted to $[-1:1]^4$ hypercube (81 sites).

• Free eigenvalues of Wilson and Brillouin operators in 4D



Free-field dispersion relations

Notation: $\hat{p}_\mu = \frac{2}{a}\sin(\frac{ap_\mu}{2})$, $\bar{p}_\mu = \frac{1}{a}\sin(ap_\mu)$, $\tilde{p}_\mu = \frac{1}{27a}\sin(ap_\mu)\prod_{\nu\neq\mu}\{\cos(ap_\nu)+2\}$

$\Rightarrow \nabla_\mu^{\rm std} = {\rm i}\bar{p}_\mu$, $\quad \triangle^{\rm std} = -\frac{4}{a^2}\sum_\mu \sin^2(\frac{ap_\mu}{2}) = \frac{2}{a^2}\sum_\mu \cos(ap_\mu) - \frac{8}{a^2} = -\sum_\mu \hat{p}_\mu^2 = -\hat{p}^2$

$\Rightarrow \nabla_\mu^{\rm iso} = {\rm i}\tilde{p}_\mu$, $\quad \triangle^{\rm bri} = \frac{4}{a^2}\prod_\mu \cos^2(\frac{ap_\mu}{2}) - \frac{4}{a^2} = \frac{1}{4a^2}\prod_\mu\{\cos(ap_\mu)+1\} - \frac{4}{a^2} \equiv -\check{p}^2$

- **Dispersion relation for Wilson operator**

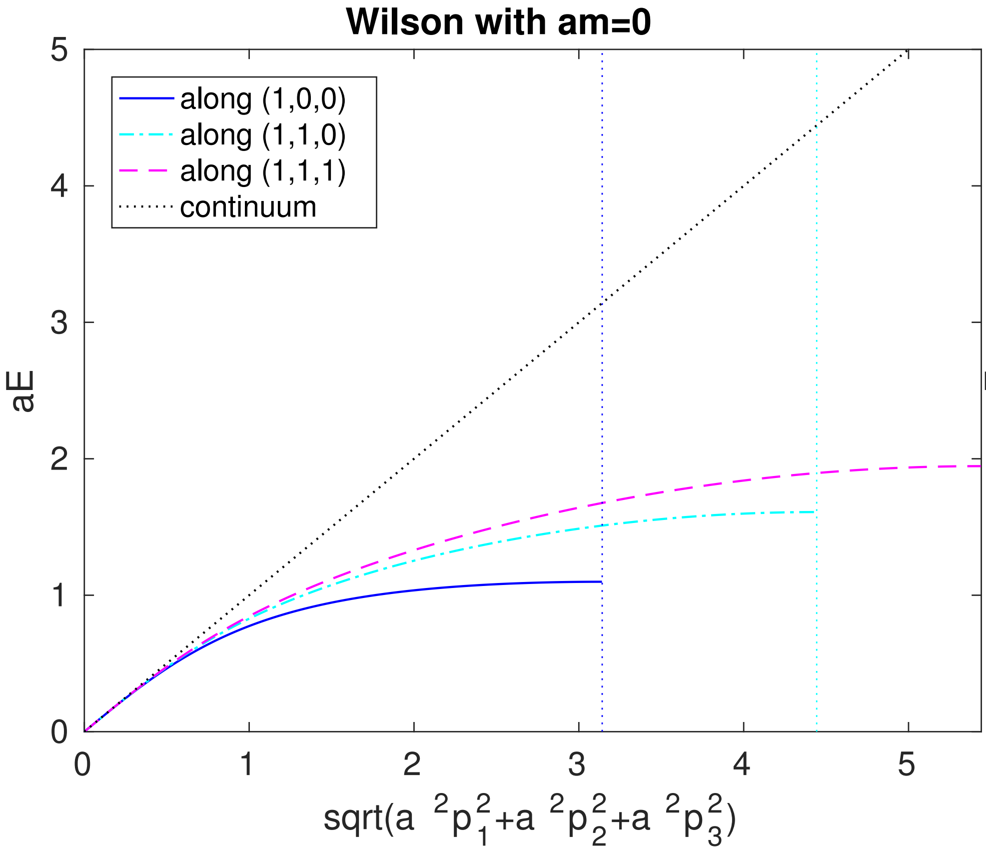
$$D_{{\rm W},m} = \nabla_\mu^{\rm std}\gamma_\mu - \frac{a}{2}\triangle^{\rm std} + m = {\rm i}\bar{p}_\mu\gamma_\mu + \frac{a}{2}\hat{p}^2 + m$$

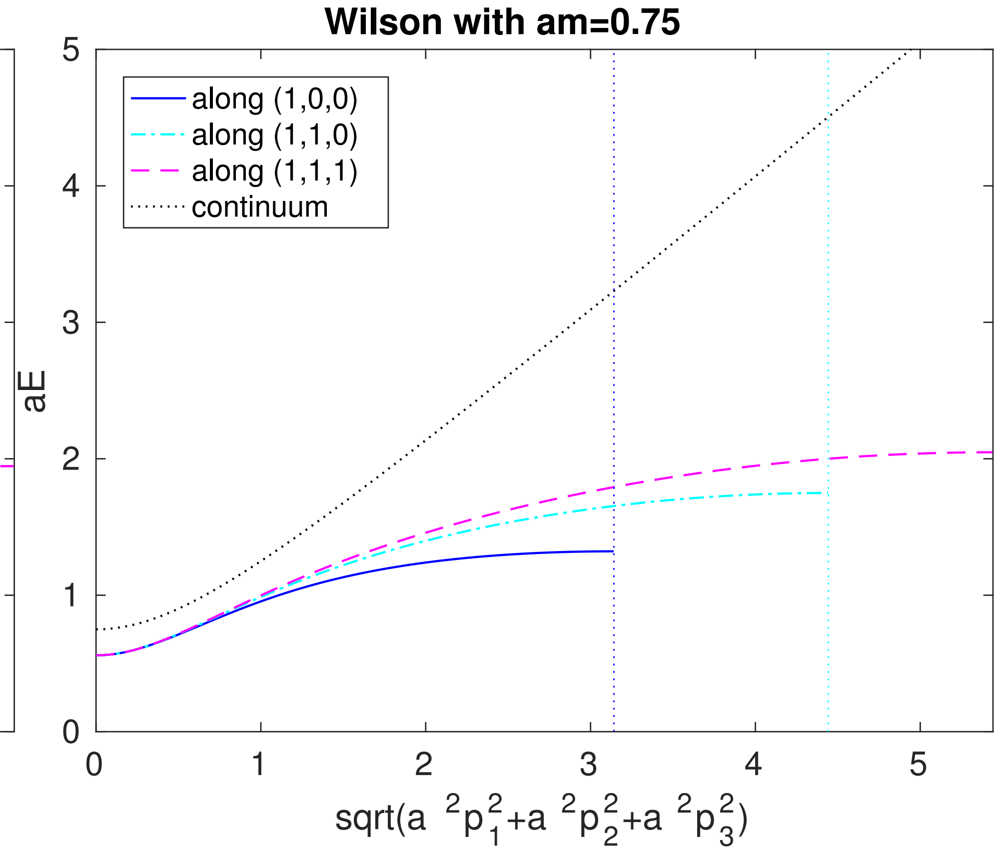
$$G_{{\rm W},m} = \frac{-{\rm i}\bar{p}_\sigma\gamma_\sigma + \frac{a}{2}\hat{p}^2 + m}{({\rm i}\bar{p}_\mu\gamma_\mu + \frac{a}{2}\hat{p}^2 + m)(-{\rm i}\bar{p}_\nu\gamma_\nu + \frac{a}{2}\hat{p}^2 + m)} = \frac{-{\rm i}\bar{p}_\sigma\gamma_\sigma + \frac{a}{2}\hat{p}^2 + m}{\bar{p}^2 + (\frac{a}{2}\hat{p}^2 + m)^2}$$

Search for zero of denominator with $p_4 \to {\rm i}E$ and $\frac{a}{2}\hat{p}^2 = -\frac{1}{a}\sum_\mu \cos(ap_\mu) + \frac{4}{a}$ yields

$$\sinh^2(aE) - \sum_i \sin^2(ap_i) = \cosh^2(aE) + 2\cosh(aE)\big[\sum_i \cos(ap_i) - 4 - am\big] + \big[...\big]^2$$

and with $\cosh^2 - \sinh^2 = 1$ this turns into a *linear equation* in $\cosh(aE)$.





- $\ominus$ strong deviation from continuum for any $a|\mathbf{p}| > 1$
- $\ominus$ strong rotational symmetry breaking for any $a|\mathbf{p}| > 1$
- $\ominus$ strong effect of $am \not\lll 1$, even at $\mathbf{p} = \mathbf{0}$

• Dispersion relation for Brillouin operator

$$D_{\mathrm{B},m} = \nabla_\mu^{\mathrm{iso}}\gamma_\mu - \frac{a}{2}\triangle^{\mathrm{bri}} + m = \mathrm{i}\tilde{p}_\mu\gamma_\mu + \frac{a}{2}\check{p}^2 + m$$

$$G_{\mathrm{B},m} = \frac{-\mathrm{i}\tilde{p}_\sigma\gamma_\sigma + \frac{a}{2}\check{p}^2 + m}{(\mathrm{i}\tilde{p}_\mu\gamma_\mu + \frac{a}{2}\check{p}^2 + m)(-\mathrm{i}\tilde{p}_\nu\gamma_\nu + \frac{a}{2}\check{p}^2 + m)} = \frac{-\mathrm{i}\tilde{p}_\sigma\gamma_\sigma + \frac{a}{2}\check{p}^2 + m}{\tilde{p}^2 + (\frac{a}{2}\check{p}^2 + m)^2}$$

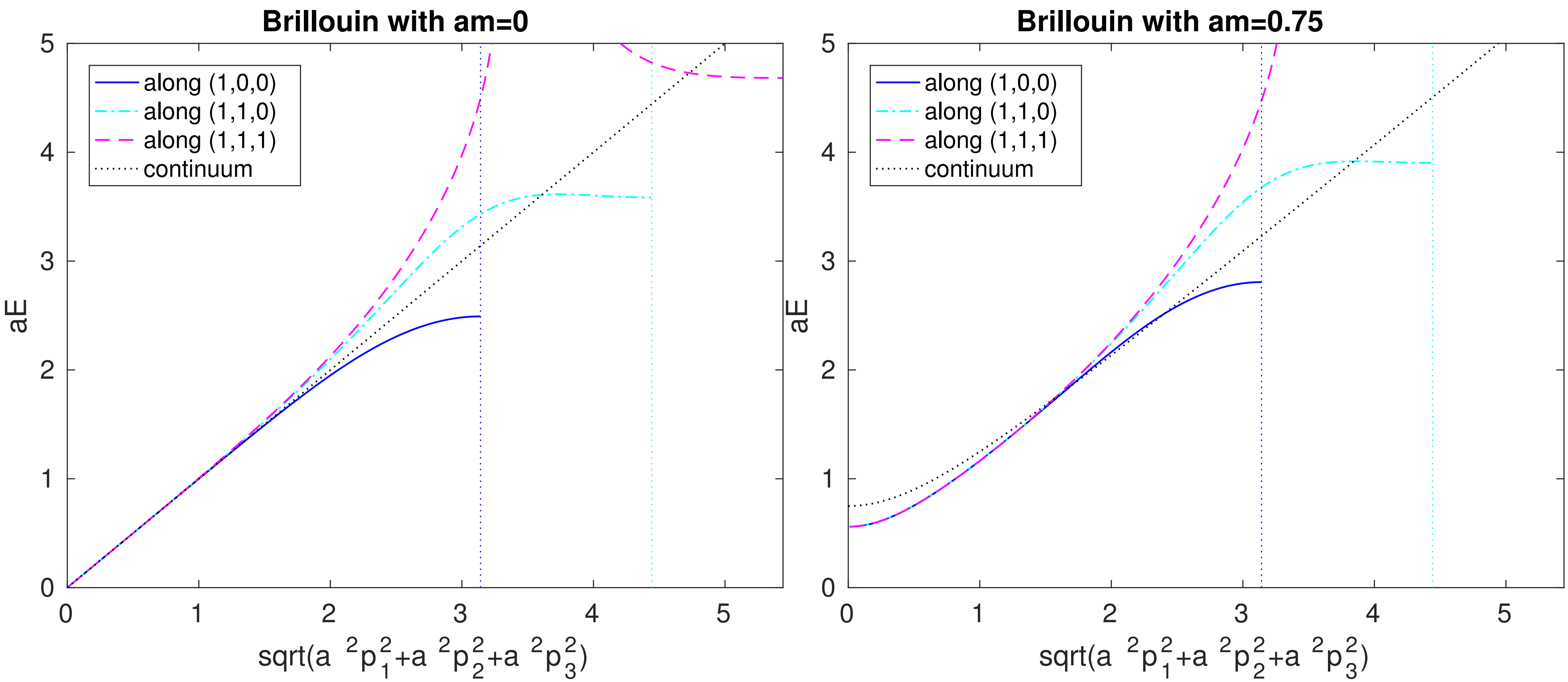
Search for zero of denominator with $p_4 \to \mathrm{i}E$ yields

$$\sum_\mu \tilde{p}_\mu^2 + \frac{1}{64a^2}\prod_\mu\{c_\mu+1\}^2 - \frac{1}{4a}\prod_\mu\{c_\mu+1\}[\frac{2}{a}+m] + [\frac{2}{a}+m]^2 = 0$$

with $\tilde{p}^2 = \frac{1}{729a^2}\sum_\mu s_\mu^2 \prod_{\nu\neq\mu}\{c_\nu+2\}^2$ this takes the form

$$\frac{1}{729a^2}\sum_i s_i^2\prod_{j\neq i}\{c_j+2\}^2\{c_4^2+4c_4+4\} + \frac{1}{729a^2}s_4^2\prod_i\{c_i+2\}^2$$
$$+\frac{1}{64a^2}\prod_i\{c_i+1\}^2\{c_4^2+2c_4+1\} - \frac{1}{4a}\prod_i\{c_i+1\}\{c_4+1\}[\frac{2}{a}+m] + [\frac{2}{a}+m]^2 = 0$$

and with $\cosh^2 - \sinh^2 = 1$ this turns into a *quadratic equation* in $\cosh(aE)$. It turns out that for any $\mathbf{p}$ explored below at most one of the solutions is real.



⊕ mild deviation from continuum up to $a|\mathbf{p}| \simeq 2$

⊕ mild rotational symmetry breaking up to $a|\mathbf{p}| \simeq 2$

⊖ strong effect of $am \not\ll 1$, especially at $\mathbf{p} = \mathbf{0}$

• Dispersion relation for overlap with Wilson kernel

$$D_{\mathrm{W},-\rho} = \mathrm{i}\bar{p}_\mu\gamma_\mu + \tfrac{a}{2}\hat{p}^2 - \tfrac{\rho}{a}$$

$$D^\dagger_{\mathrm{W},-\rho}D_{\mathrm{W},-\rho} = (-\mathrm{i}\bar{p}_\mu\gamma_\mu + \tfrac{a}{2}\hat{p}^2 - \tfrac{\rho}{a})(\mathrm{i}\bar{p}_\nu\gamma_\nu + \tfrac{a}{2}\hat{p}^2 - \tfrac{\rho}{a}) = \bar{p}^2 + (\tfrac{a}{2}\hat{p}^2 - \tfrac{\rho}{a})^2$$

$$D_{\mathrm{NW},m} = \big(1 - \tfrac{am}{2\rho}\big)D_{\mathrm{NW}} + m \quad \text{with} \quad D_{\mathrm{NW}} = \tfrac{\rho}{a}\big\{1 + D_{\mathrm{W},-\rho}[D^\dagger_{\mathrm{W},-\rho}D_{\mathrm{W},-\rho}]^{-1/2}\big\}$$

$$D_{\mathrm{NW},m} = \underbrace{\big(\tfrac{\rho}{a} + \tfrac{m}{2}\big)}_{\equiv c} + \underbrace{\big(\tfrac{\rho}{a} - \tfrac{m}{2}\big)}_{\equiv d}\big(\mathrm{i}\bar{p}_\mu\gamma_\mu + \tfrac{a}{2}\hat{p}^2 - \tfrac{\rho}{a}\big)\big[\bar{p}^2 + (\tfrac{a}{2}\hat{p}^2 - \tfrac{\rho}{a})^2\big]^{-1/2}$$

$$G_{\mathrm{NW},m} = \frac{c+d[\bar{p}^2+(\frac{a}{2}\hat{p}^2-\frac{\rho}{a})^2]^{-1/2}(-\mathrm{i}\bar{p}_\sigma\gamma_\sigma+\frac{a}{2}\hat{p}^2-\frac{\rho}{a})}{\Big\{c+d[\bar{p}^2+(\frac{a}{2}\hat{p}^2-\frac{\rho}{a})^2]^{-1/2}(-\mathrm{i}\bar{p}_\mu\gamma_\mu+\frac{a}{2}\hat{p}^2-\frac{\rho}{a})\Big\}\Big\{c+d(\mathrm{i}\bar{p}_\nu\gamma_\nu+\frac{a}{2}\hat{p}^2-\frac{\rho}{a})[\bar{p}^2+(\frac{a}{2}\hat{p}^2-\frac{\rho}{a})^2]^{-1/2}\Big\}}$$

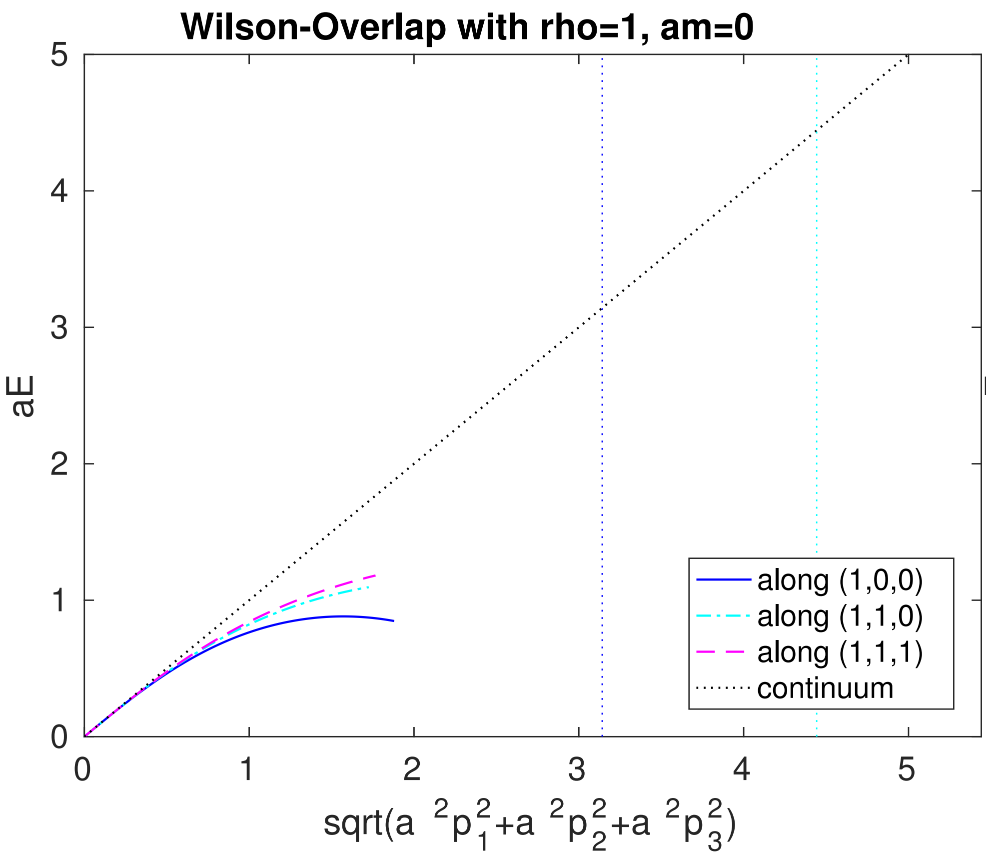
$$= \frac{c+d[\bar{p}^2+(\frac{a}{2}\hat{p}^2-\frac{\rho}{a})^2]^{-1/2}(-\mathrm{i}\bar{p}_\sigma\gamma_\sigma+\frac{a}{2}\hat{p}^2-\frac{\rho}{a})}{c^2+2cd(\frac{a}{2}\hat{p}^2-\frac{\rho}{a})[\ldots]^{-1/2}+d^2[\ldots]^{-1/2}[\ldots][\ldots]^{-1/2}}$$

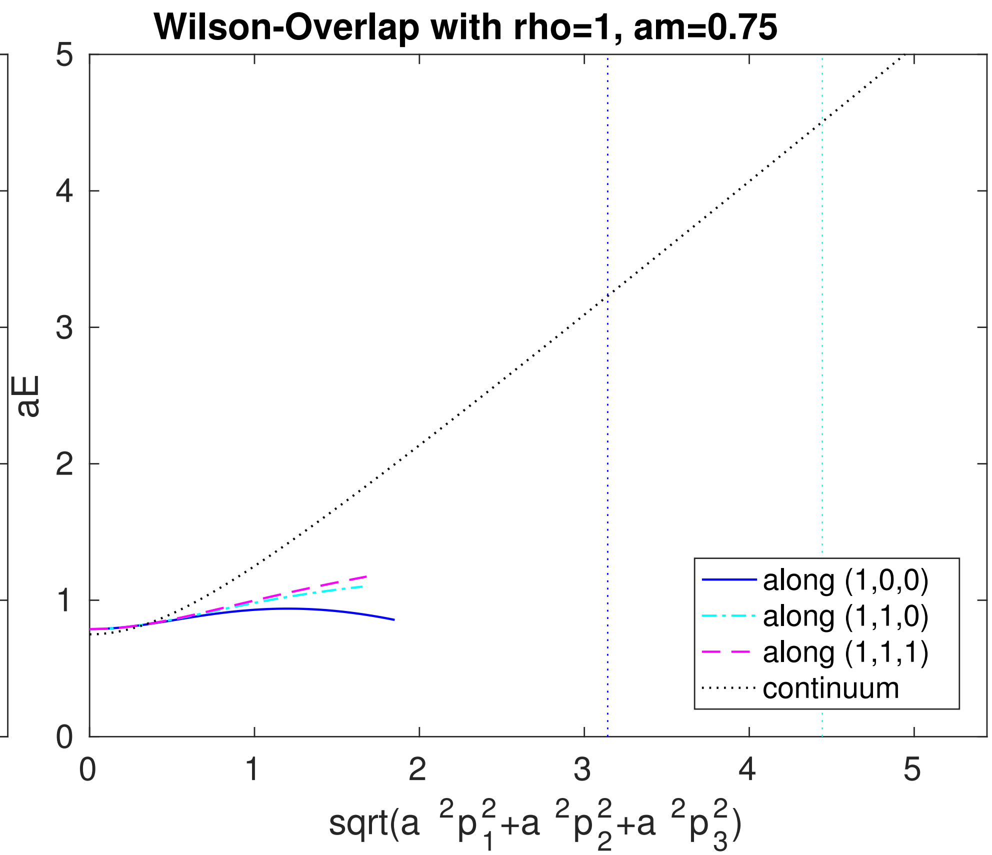
End up searching for zero in $c^2 + 2cd(\frac{a}{2}\hat{p}^2 - \frac{\rho}{a})[\bar{p}^2 + (\frac{a}{2}\hat{p}^2 - \frac{\rho}{a})^2]^{-1/2} + d^2 = 0$ with

$$\bar{p}^2 = \tfrac{1}{a^2}\textstyle\sum_\mu s_\mu^2 \qquad \text{with} \qquad s_\mu \equiv \sin(ap_\mu)$$

$$\hat{p}^2 = \tfrac{8}{a^2} - \tfrac{2}{a^2}\textstyle\sum_\mu c_\mu \qquad \text{with} \qquad c_\mu \equiv \cos(ap_\mu)$$

and the inverse sqare root renders this a *transcendental equation* in $\cosh(aE)$.





- ⊖ strong deviation from continuum for any $a|\mathbf{p}| > 1$
- ⊖ strong rotational symmetry breaking for any $a|\mathbf{p}| > 1$
- ⊕ mild effect of $am \not\ll 1$, at least at $\mathbf{p} = \mathbf{0}$

- **Dispersion relation for overlap with Brillouin kernel**

$$D_{\mathrm{B},-\rho} = \mathrm{i}\tilde{p}_\mu\gamma_\mu + \tfrac{a}{2}\check{p}^2 - \tfrac{\rho}{a}$$

$$D^\dagger_{\mathrm{B},-\rho} D_{\mathrm{B},-\rho} = (-\mathrm{i}\tilde{p}_\mu\gamma_\mu + \tfrac{a}{2}\check{p}^2 - \tfrac{\rho}{a})(\mathrm{i}\tilde{p}_\nu\gamma_\nu + \tfrac{a}{2}\check{p}^2 - \tfrac{\rho}{a}) = \tilde{p}^2 + (\tfrac{a}{2}\check{p}^2 - \tfrac{\rho}{a})^2$$

$$D_{\mathrm{NB},m} = \big(1-\tfrac{am}{2\rho}\big) D_{\mathrm{NB}} + m \quad \text{with} \quad D_{\mathrm{NB}} = \tfrac{\rho}{a}\big\{1 + D_{\mathrm{B},-\rho}[D^\dagger_{\mathrm{B},-\rho} D_{\mathrm{B},-\rho}]^{-1/2}\big\}$$

$$D_{\mathrm{NB},m} = \underbrace{\big(\tfrac{\rho}{a}+\tfrac{m}{2}\big)}_{\equiv c} + \underbrace{\big(\tfrac{\rho}{a}-\tfrac{m}{2}\big)}_{\equiv d} \big(\mathrm{i}\tilde{p}_\mu\gamma_\mu + \tfrac{a}{2}\check{p}^2 - \tfrac{\rho}{a}\big)\big[\tilde{p}^2 + (\tfrac{a}{2}\check{p}^2 - \tfrac{\rho}{a})^2\big]^{-1/2}$$

$$G_{\mathrm{NB},m} = \frac{c+d[\tilde{p}^2+(\frac{a}{2}\check{p}^2-\frac{\rho}{a})^2]^{-1/2}(-\mathrm{i}\tilde{p}_\sigma\gamma_\sigma+\frac{a}{2}\check{p}^2-\frac{\rho}{a})}{\Big\{c+d[\tilde{p}^2+(\frac{a}{2}\check{p}^2-\frac{\rho}{a})^2]^{-1/2}(-\mathrm{i}\tilde{p}_\mu\gamma_\mu+\frac{a}{2}\check{p}^2-\frac{\rho}{a})\Big\}\Big\{c+d(\mathrm{i}\tilde{p}_\nu\gamma_\nu+\frac{a}{2}\check{p}^2-\frac{\rho}{a})[\tilde{p}^2+(\frac{a}{2}\check{p}^2-\frac{\rho}{a})^2]^{-1/2}\Big\}}$$

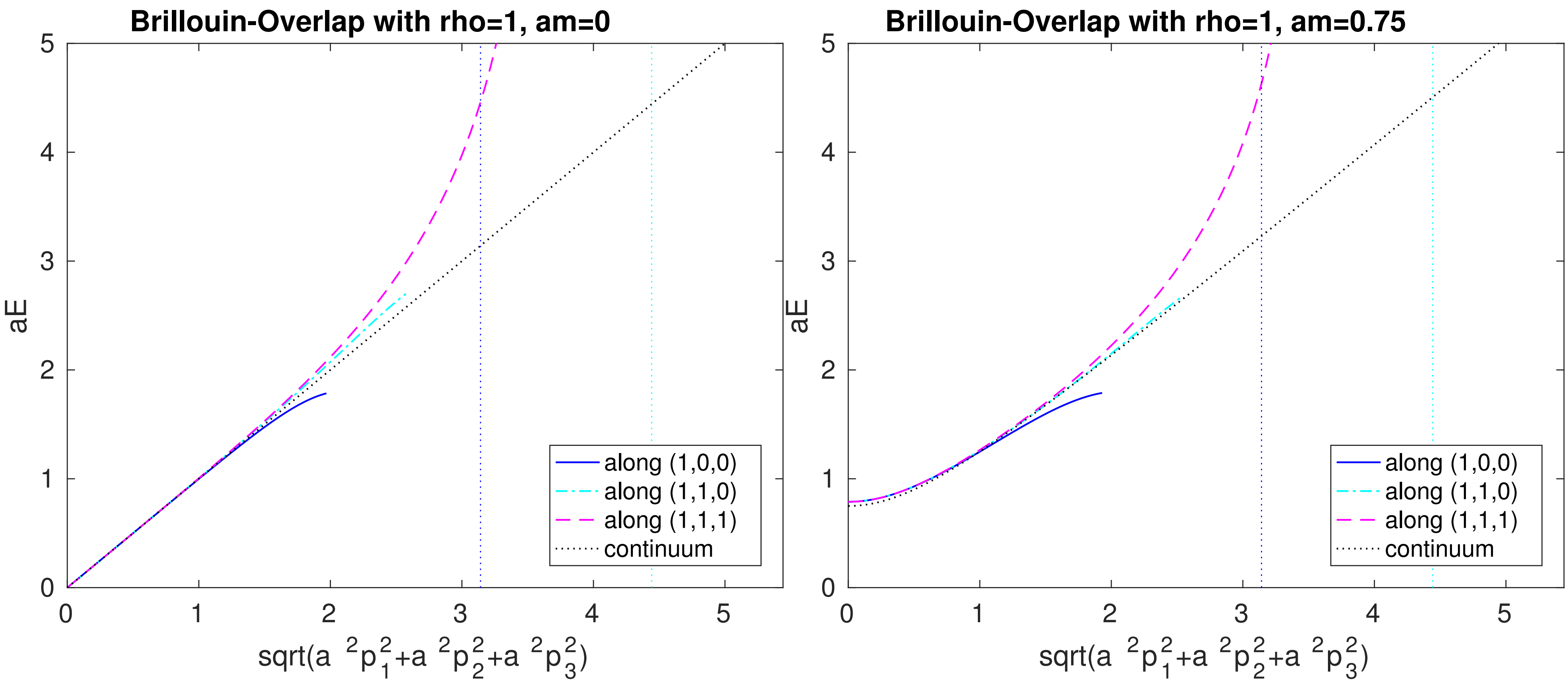
$$= \frac{c+d[\tilde{p}^2+(\frac{a}{2}\check{p}^2-\frac{\rho}{a})^2]^{-1/2}(-\mathrm{i}\tilde{p}_\sigma\gamma_\sigma+\frac{a}{2}\check{p}^2-\frac{\rho}{a})}{c^2+2cd(\frac{a}{2}\check{p}^2-\frac{\rho}{a})[\ldots]^{-1/2}+d^2[\ldots]^{-1/2}[\ldots][\ldots]^{-1/2}}$$

End up searching for zero in $c^2 + 2cd(\tfrac{a}{2}\check{p}^2 - \tfrac{\rho}{a})[\tilde{p}^2 + (\tfrac{a}{2}\check{p}^2 - \tfrac{\rho}{a})^2]^{-1/2} + d^2 = 0$ with

$$\tilde{p}^2 = \tfrac{1}{729a^2}\textstyle\sum_\mu s_\mu^2 \prod_{\nu\neq\mu}\{c_\nu+2\}^2$$

$$\check{p}^2 = \tfrac{4}{a^2} - \tfrac{1}{4a^2}\textstyle\prod_\mu\{c_\mu+1\}$$

and the inverse sqare root renders this a *transcendental equation* in $\cosh(aE)$.

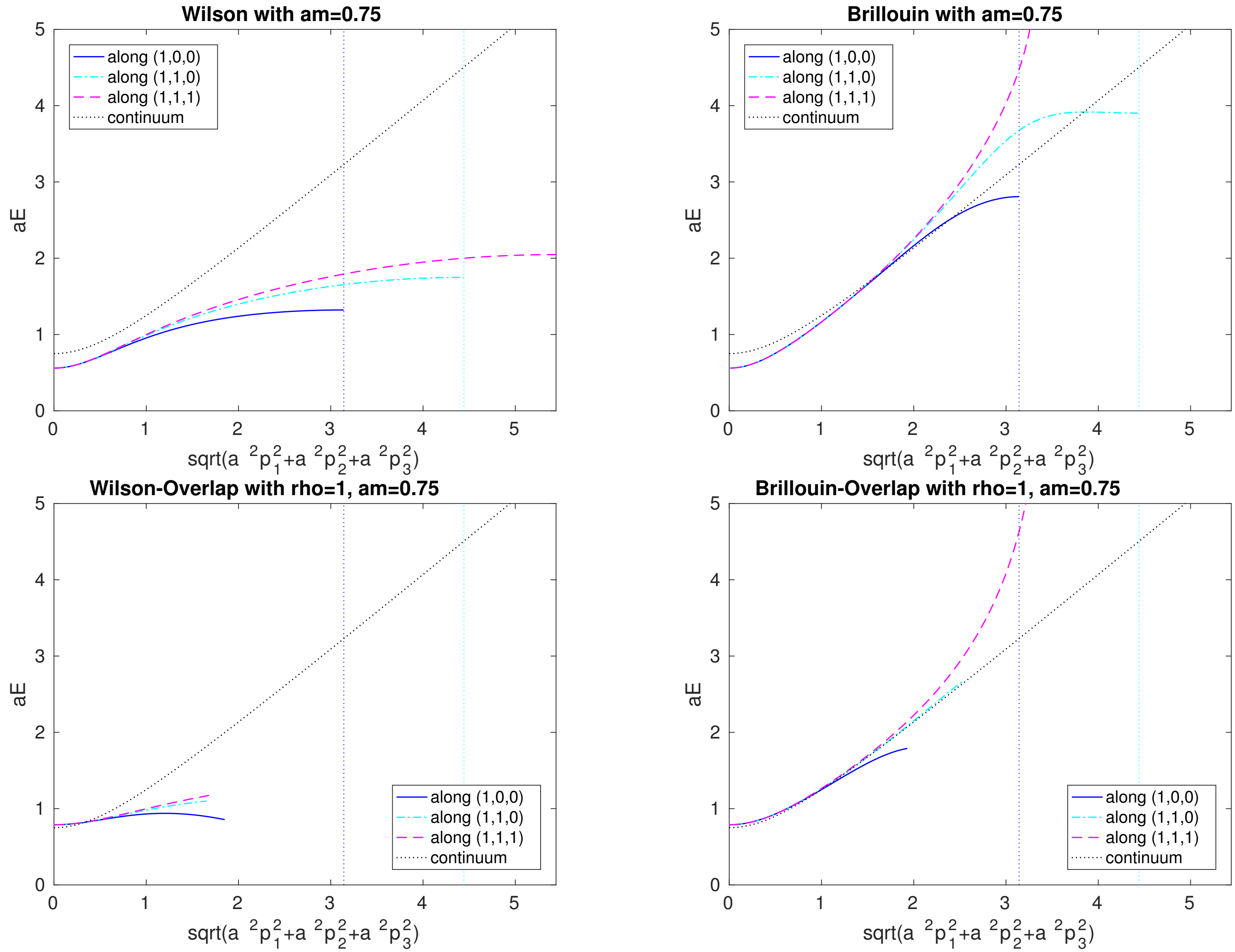


⊕ mild deviation from continuum up to $a|\mathbf{p}| \simeq 2$

⊕ mild rotational symmmetry breaking up to $a|\mathbf{p}| \simeq 2$

⊕ mild effect of $am \not\lll 1$ as long as $|\mathbf{p}|$ not loo large

• Interim summary on dispersion relations



⟶ Brillouin kernel improves DR w.r.t. momenta, but *not* w.r.t. quark masses

⟶ Overlap recipe improves DR w.r.t. quark masses, but *not* w.r.t. momenta

⟹ Brillouin-Overlap shows reasonable DR w.r.t. both masses and momenta

Brillouin kernel: Implementation details

"*prohibitively expensive*" or "*just expensive*" ?

- **Overall smearing**

Use same smearing in covariant derivative and Laplacian and improvement term.
Practical solution: smeared gauge field copy, i.e. create `V(Nc,Nc,4,Nx,Ny,Nz,Nt)` alongside original U_ρ, and evaluate Wilson/Brillouin operator on V_ρ.
For $c_\mathrm{SW}\neq 0$ precompute $F_{\mu\nu}$ from V_ρ, i.e. create `F(Nc,Nc,6,Nx,Ny,Nz,Nt)`.

- **Gauging strategy**

To maintain γ_5-hermiticity: sum over all paths of equal length [1 path for 1-hop, 2 paths for 2-hop, 6 paths for 3-hop, 24 paths for 4-hop, average or SU(3)-project].
Practical solution: off-axis-link precomputed, i.e. create `W(Nc,Nc,40,Nx,Ny,Nz,Nt)`.

- **Bandwidth saturation**

Practical solution: multiple-vector strategy, i.e. use `vec(Nc,4,Nvec,Nx*Ny*Nz*Nt)`

- **Routine `app_bril_sp/dp`**

`GK`: OpenMP+MPI routine in C++, see https://github.com/g-koutsou/qpb.
`SD`: OpenMP-only routine in F2008, discussed on following two slides.

```
!$OMP PARALLEL DO DEFAULT(private) FIRSTPRIVATE(Nx,Ny,Nz,Nt,Nvec,mass) SHARED(old,new,W) SCHEDULE(static)
do l=1,Nt
do k=1,Nz
do j=1,Ny
do i=1,Nx
   n=(((l-1)*Nz+(k-1))*Ny+(j-1))*Nx+i
   site(:,:,:)=mass*old(:,:,:,n) !!! note: site is Nc*4*Nvec
   !!! visit all 81 sites within hypercube (distances 0 to 4 in taxi-driver metric)
   do go_l=-1,1; lsh=modulo(l+go_l-1,Nt)+1
   do go_k=-1,1; ksh=modulo(k+go_k-1,Nz)+1
   do go_j=-1,1; jsh=modulo(j+go_j-1,Ny)+1
   do go_i=-1,1; ish=modulo(i+go_i-1,Nx)+1
      nsh=(((lsh-1)*Nz+(ksh-1))*Ny+(jsh-1))*Nx+ish
      dir=(go_l+1)*27+(go_k+1)*9+(go_j+1)*3+go_i+2
      select case(dir)
         case(01:40); tmp=W(:,:,dir,i,j,k,l)
         case(   41); tmp=color_eye() !!! note: yields Nc*Nc identity matrix
         case(42:81); tmp=conjg(transpose(W(:,:,82-dir,ish,jsh,ksh,lsh)))
      end select
      select case(abs(go_i)+abs(go_j)+abs(go_k)+abs(go_l))
         case(0); fac_i=            0.0 ; fac_j=     0.0 ; fac_k=     0.0 ; fac_l=     0.0 ; fac=(-240.0/128.0)
         case(1); fac_i=go_i*(64.0/432.0); fac_j=go_j*(...); fac_k=go_k*(...); fac_l=go_l*(...); fac=(   8.0/128.0)
         case(2); fac_i=go_i*(16.0/432.0); fac_j=go_j*(...); fac_k=go_k*(...); fac_l=go_l*(...); fac=(   4.0/128.0)
         case(3); fac_i=go_i*( 4.0/432.0); fac_j=go_j*(...); fac_k=go_k*(...); fac_l=go_l*(...); fac=(   2.0/128.0)
         case(4); fac_i=go_i*( 1.0/432.0); fac_j=go_j*(...); fac_k=go_k*(...); fac_l=go_l*(...); fac=(   1.0/128.0)
      end select
      do idx=1,Nvec
         tmq=matmul(tmp,old(:,:,idx,nsh)) !!! note: tmq is Nc*4
         !!! add terms proportional to isotropic derivative operators
         if (go_i.ne.0) then
            site(:,1,idx)=site(:,1,idx)-cmplx(0.0,fac_i)*tmq(:,4) !!! transpose(gamma1)=   0   0   0   i
            site(:,2,idx)=site(:,2,idx)-cmplx(0.0,fac_i)*tmq(:,3) !!!                      0   0   i   0
            site(:,3,idx)=site(:,3,idx)+cmplx(0.0,fac_i)*tmq(:,2) !!!                      0  -i   0   0
            site(:,4,idx)=site(:,4,idx)+cmplx(0.0,fac_i)*tmq(:,1) !!!                     -i   0   0   0
         end if
```

```
        if (go_j.ne.0) then
           site(:,1,idx)=site(:,1,idx)-fac_j*tmq(:,4)            !!! transpose(gamma2)=    0     0     0    -1
           site(:,2,idx)=site(:,2,idx)+fac_j*tmq(:,3)            !!!                       0     0     1     0
           site(:,3,idx)=site(:,3,idx)+fac_j*tmq(:,2)            !!!                       0     1     0     0
           site(:,4,idx)=site(:,4,idx)-fac_j*tmq(:,1)            !!!                      -1     0     0     0
        end if
        if (go_k.ne.0) then
           site(:,1,idx)=site(:,1,idx)-cmplx(0.0,fac_k)*tmq(:,3) !!! transpose(gamma3)=    0     0     i     0
           site(:,2,idx)=site(:,2,idx)+cmplx(0.0,fac_k)*tmq(:,4) !!!                       0     0     0    -i
           site(:,3,idx)=site(:,3,idx)+cmplx(0.0,fac_k)*tmq(:,1) !!!                      -i     0     0     0
           site(:,4,idx)=site(:,4,idx)-cmplx(0.0,fac_k)*tmq(:,2) !!!                       0     i     0     0
        end if
        if (go_l.ne.0) then
           site(:,1,idx)=site(:,1,idx)+fac_l*tmq(:,3)            !!! transpose(gamma4)=    0     0     1     0
           site(:,2,idx)=site(:,2,idx)+fac_l*tmq(:,4)            !!!                       0     0     0     1
           site(:,3,idx)=site(:,3,idx)+fac_l*tmq(:,1)            !!!                       1     0     0     0
           site(:,4,idx)=site(:,4,idx)+fac_l*tmq(:,2)            !!!                       0     1     0     0
        end if
        !!! subtract 1/2 times brillouin laplacian
        site(:,:,idx)=site(:,:,idx)-fac*tmq !!! note: factor 0.5 excluded here but shuffled into "fac"
     end do ! idx=1,Nvec
   end do ! go_i=-1,1
   end do ! go_j=-1,1
   end do ! go_k=-1,1
   end do ! go_l=-1,1
   !!! plug everything into new vector
   do idx=1,Nvec
      new(:,:,idx,n)=site(:,:,idx)
   end do ! idx=1,Nvec
end do ! i=1,Nx
end do ! j=1,Ny
end do ! k=1,Nz
end do ! l=1,Nt
!$OMP END PARALLEL DO
```

- **Timings on $24^3 \times 48$ lattice**

```
ifort [...] -openmp -O2 -xavx -opt-mem-bandwidth1 -c modulename.f90
```

Matrix-vector multiplication in sp/dp on 4-core CPU [sec/vec]:

	Wilson	Brillouin	ratio
$c_{\mathrm{SW}} = 0$	0.033	0.611	18.5
$c_{\mathrm{SW}} = 1$	0.059	0.636	10.8

BiCGstab unpreconditioned inversion in mp/dp [sec/rhs]:

	Wilson	Brillouin	ratio
$c_{\mathrm{SW}} = 0$	18.6	88.7	4.8
$c_{\mathrm{SW}} = 1$	40.9	179.7	4.4
iterations	45+28+35=108	57	—
iterations	74+53+57=184	112	—

($c_{\mathrm{SW}}\!=\!0$ with $am\!=\!0.1$ and 7 stout steps results in $aM_\pi \simeq 0.65$ in both cases)
($c_{\mathrm{SW}}\!=\!1$ with $am\!=\!0.1$ and 7 stout steps results in $aM_\pi \simeq 0.45$ in both cases)

Summary: matrix-vector is **10-20** times more expensive with Brillouin action, but a factor **2-4** comes back (reduced iteration count, fixed linalg), hence overall factor **5**.

• Wilson flop count

For matrix-times-vector operation we must (per site):
(i) spin-project (from 4 to 2) for each direction $\longrightarrow 12 \cdot 8 = 96$ flops
(ii) SU(3)-multiply (spin-reduced) for each direction $\longrightarrow 6 \cdot 22 \cdot 8 = 1056$ flops
(iii) accumulate 8+mass 1 contributions to out-spinor $\longrightarrow 24 \cdot 9 = 216$ flops

All together 1368 flops per site and $24^3 \times 48 = 663552$ sites. Performance in simple OpenMP implementation is $1368 \cdot 663552/0.033$ flops per second or 27.5 Gflops.

• Brillouin flop count

For matrix-times-vector operation we must (per site):
(i) SU(3)-multiply (spin-full) for each direction $\longrightarrow 12 \cdot 22 \cdot 80 = 21120$ flops
(ii) multiply with fac_i/.../fac: $24 \cdot (54{+}54{+}54{+}54{+}81) = 7128$ flops
(iii) accumulate 80+mass 1 contributions to out-spinor $\longrightarrow 24 \cdot 81 = 1944$ flops

All together 30192 flops per site and $24^3 \times 48 = 663552$ sites. Performance in simple OpenMP implementation is $30192 \cdot 663552/0.611$ flops per second or 32.8 Gflops.

• Summary

The Brillouin-to-Wilson ratio of flops is **22.1**, while observed timing ratio was **18.5**.

- **Wilson memory traffic**

For matrix-times-vector operation we must (per site, with #rhs=1):
(i) read one sp-spinor for each direction $\longrightarrow 24 \cdot 9 = 216$ floats
(ii) read one sp-gauge matrix `V` per direction $\longrightarrow 18 \cdot 8 = 144$ floats
(iii) write one sp-spinor $\longrightarrow 24$ floats

All together 384 floats of traffic per site, i.e. 1536 bytes in sp (1.12 bytes/flop).

- **Brillouin memory traffic**

For matrix-times-vector operation we must (per site, with #rhs=1):
(i) read one sp-spinor for each direction $\longrightarrow 24 \cdot 81 = 1944$ floats
(ii) read one sp-gauge matrix `W` per direction $\longrightarrow 18 \cdot 80 = 1440$ floats
(iii) write one sp-spinor $\longrightarrow 24$ floats

All together 3408 floats of traffic per site, i.e. 13632 bytes in sp (0.45 bytes/flop).

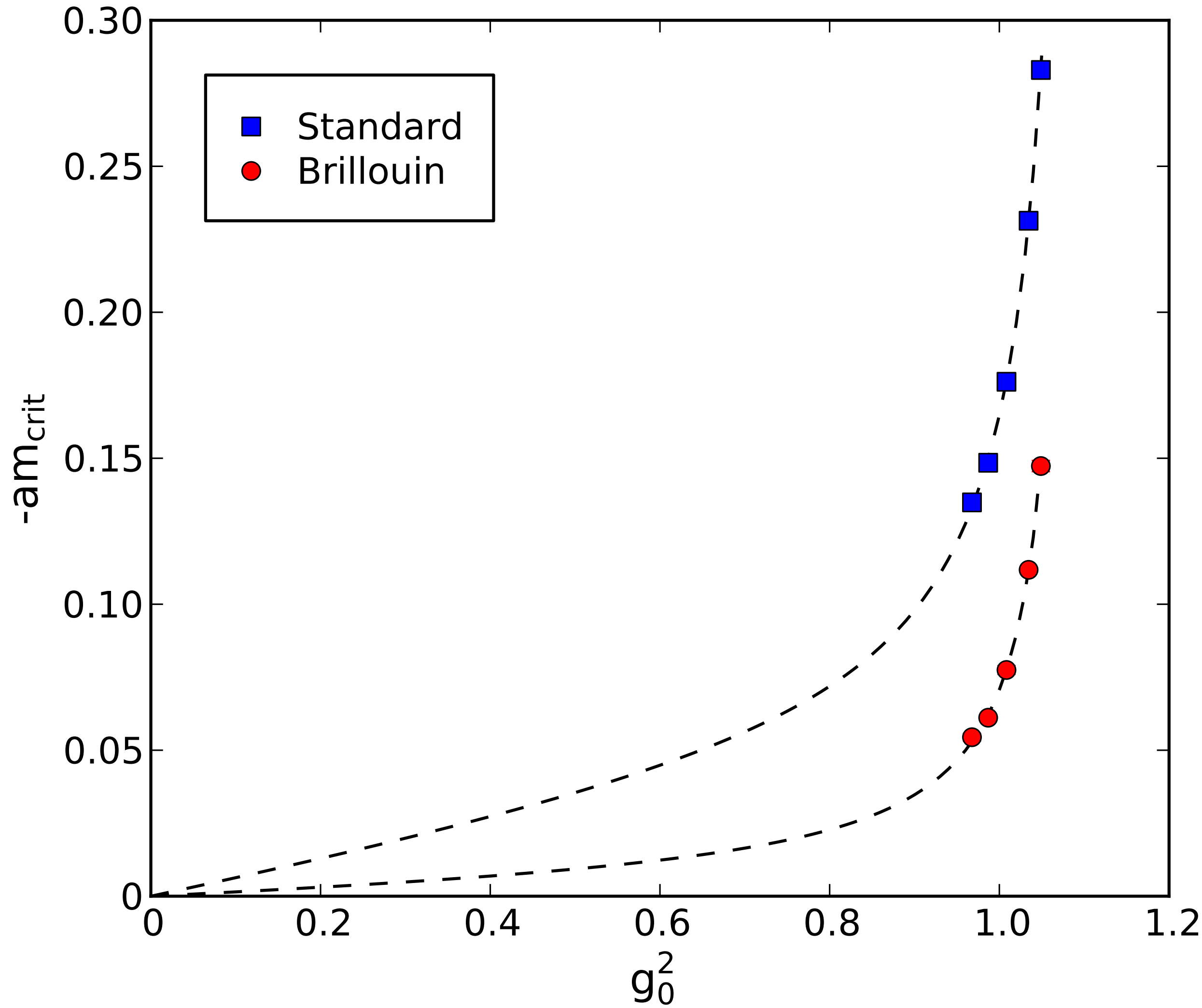
- **Summary/Comments**

The Brillouin-to-Wilson ratio of memory traffic is 3408/384=**8.9** or 2088/252=**8.9**.
With 12 rhs W changes to 216+12+24=252 floats from/to memory per site.
With 12 rhs B changes to 1944+120+24=2088 floats from/to memory per site.
Worst case scenario assumed, i.e. everything to be read afresh, since no cache.

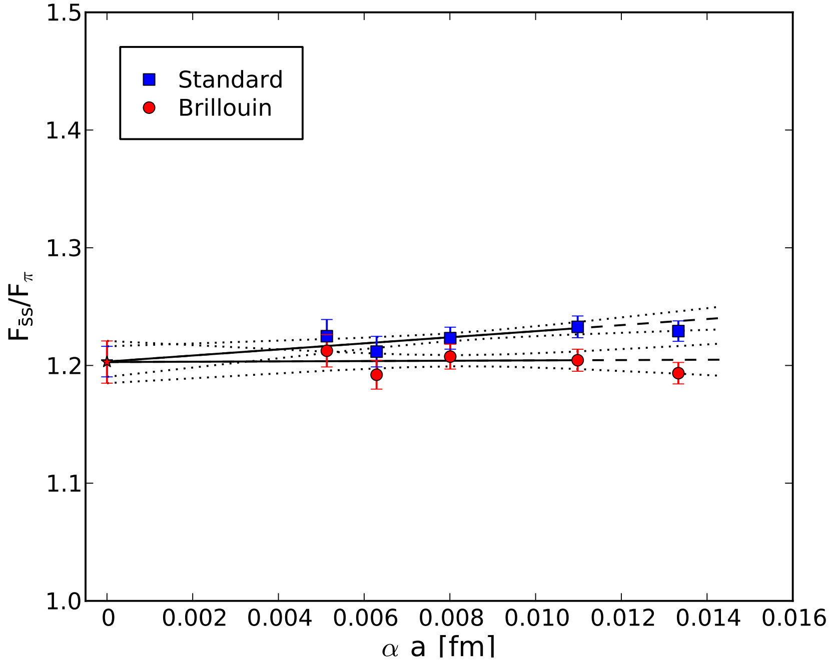
Brillouin kernel: Spectroscopy tests

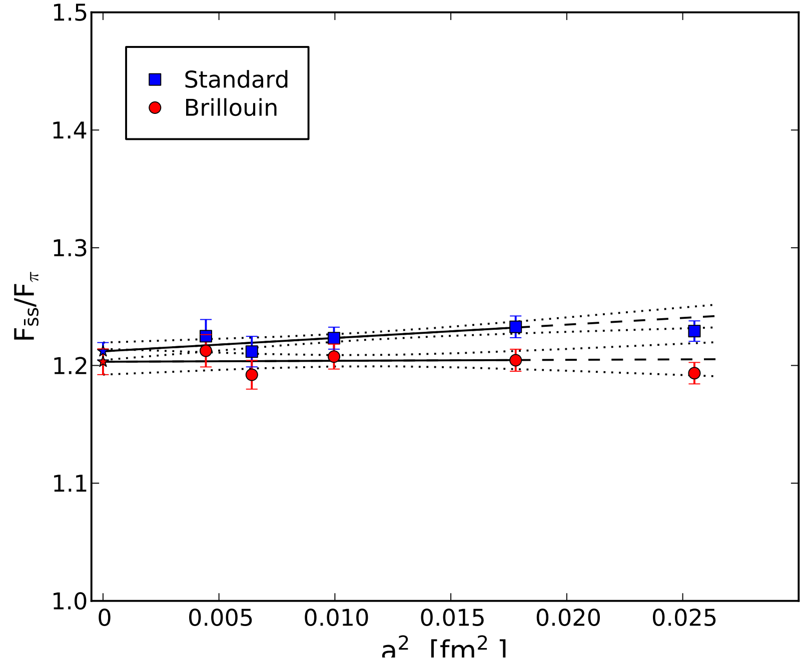
Throughout this section: 1 APE step with $\alpha = 0.72$ in $\nabla^{\rm iso}, \triangle^{\rm bri}, F_{\mu\nu}$ and $c_{\rm SW} = 1$

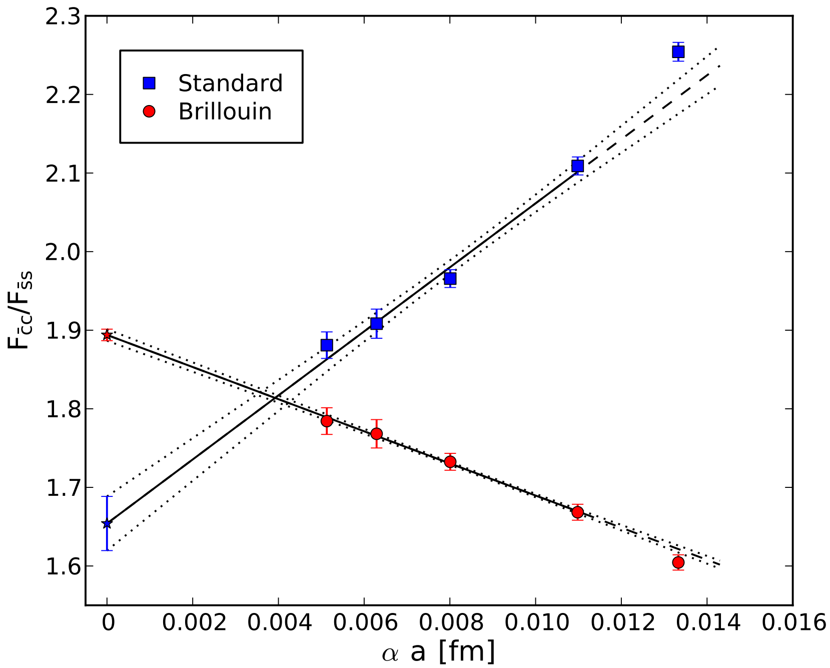
- **Additive mass renormalization**

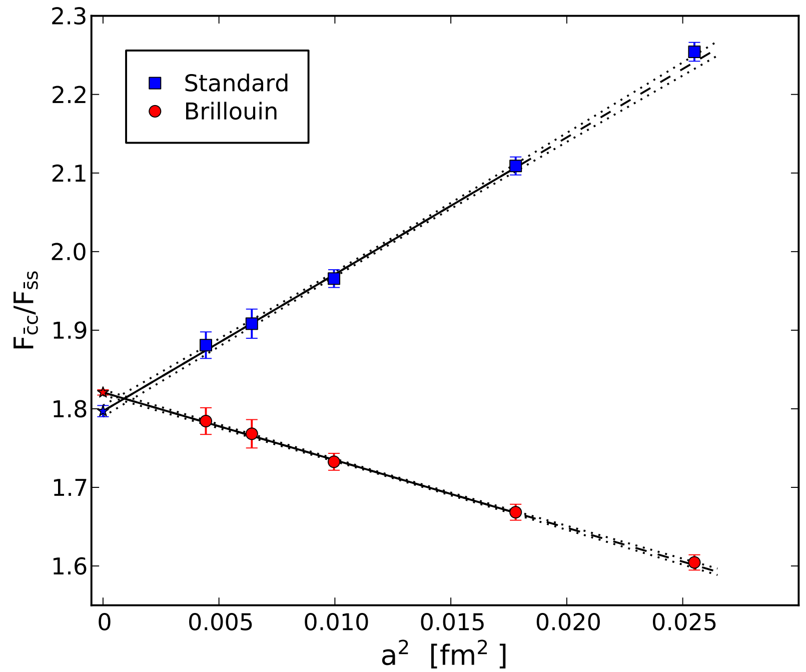


• Ratios of decay constants

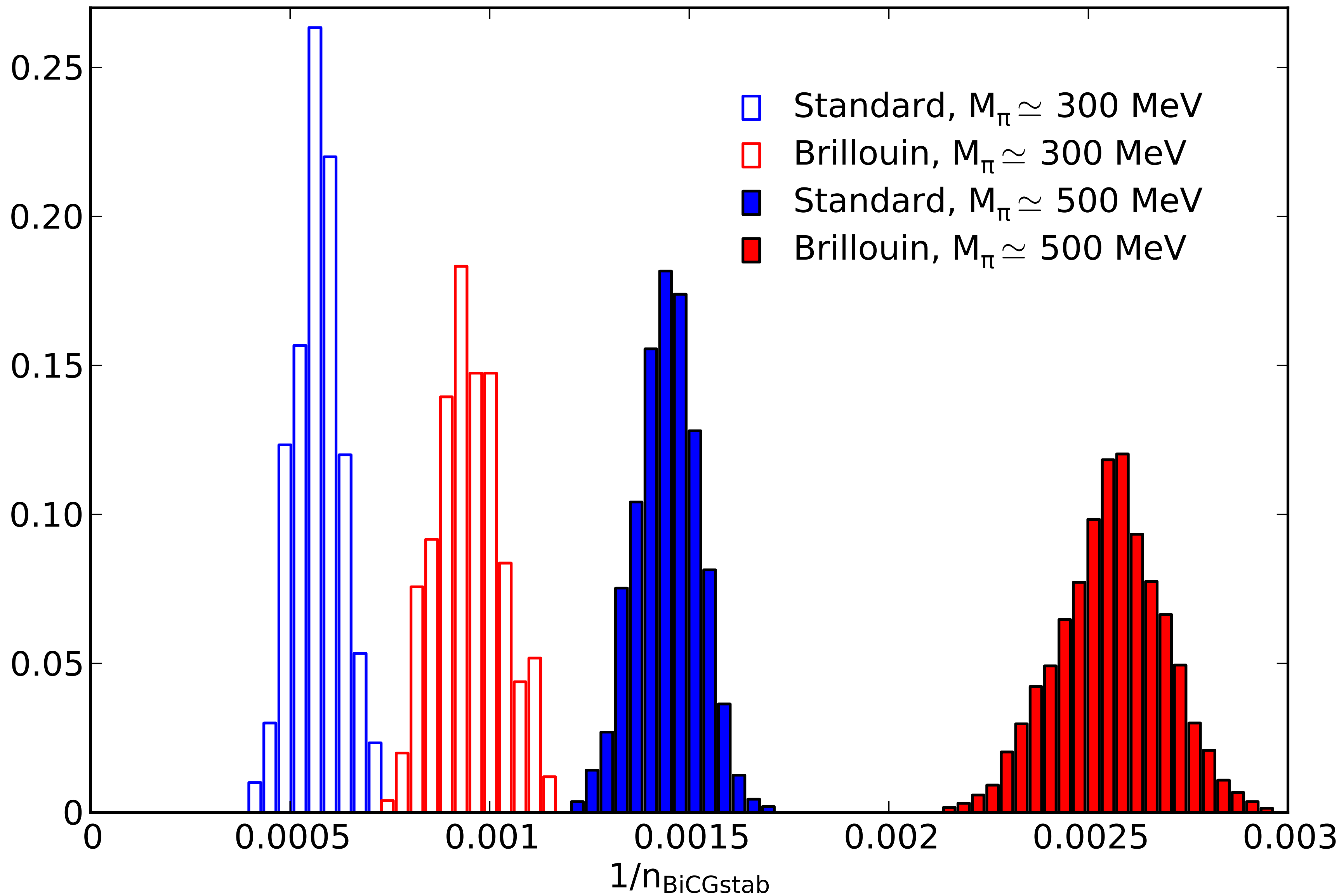




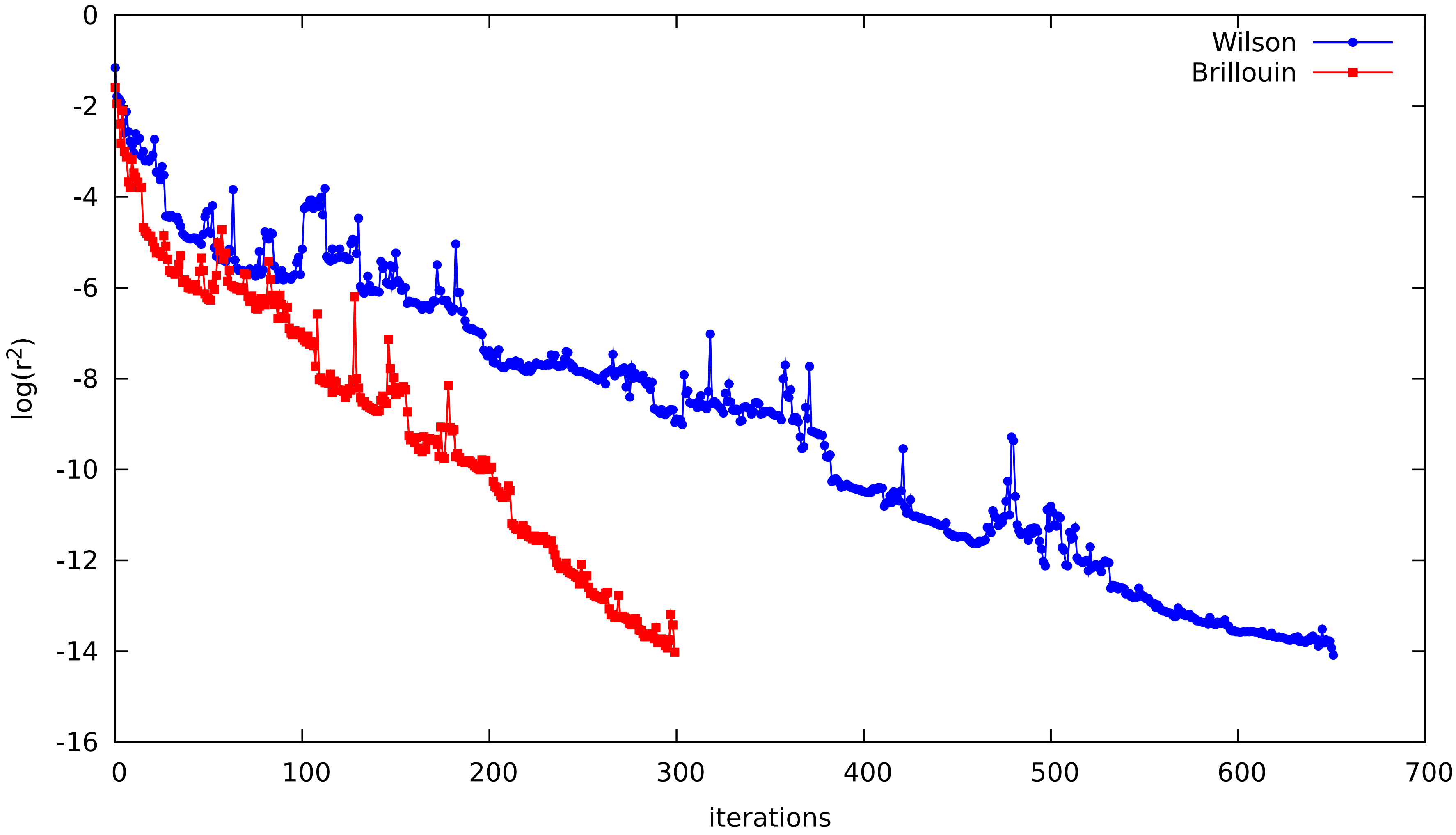




- **Inverse iteration count**



• BiCGstab convergence history

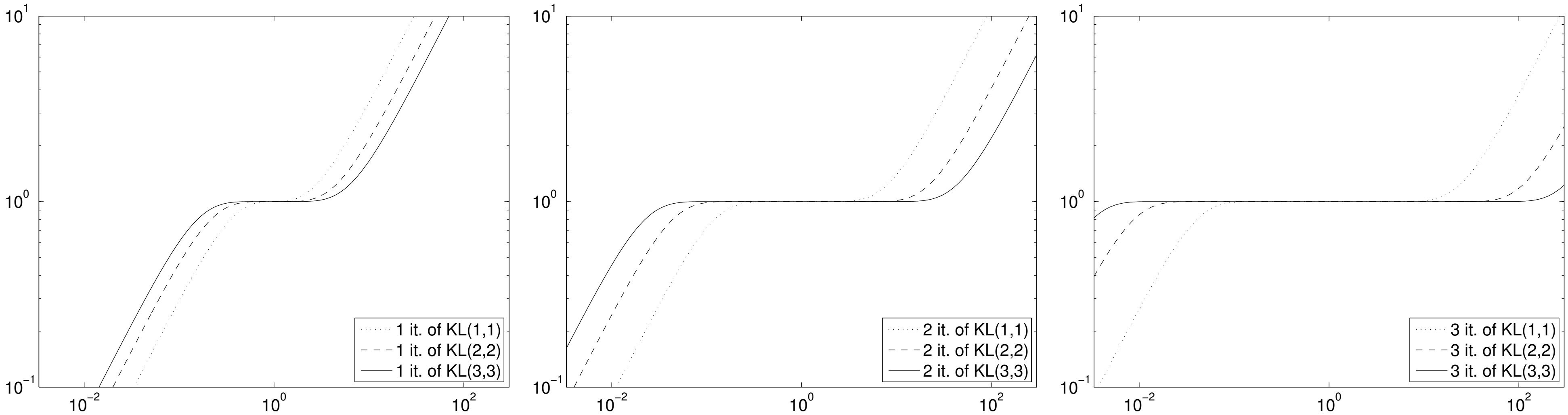


Brillouin overlap: Implementation details

$$D_{\mathrm{overlap},0} = \frac{\rho}{a}\{1+\gamma_5 \mathrm{sign}(\gamma_5 D_{\mathrm{kernel},-\rho/a})\}$$
$$= \frac{\rho}{a}\{1+D_{\mathrm{ker},-\rho/a}(D^\dagger_{\mathrm{ker},-\rho/a}D_{\mathrm{ker},-\rho/a})^{-1/2}\}$$

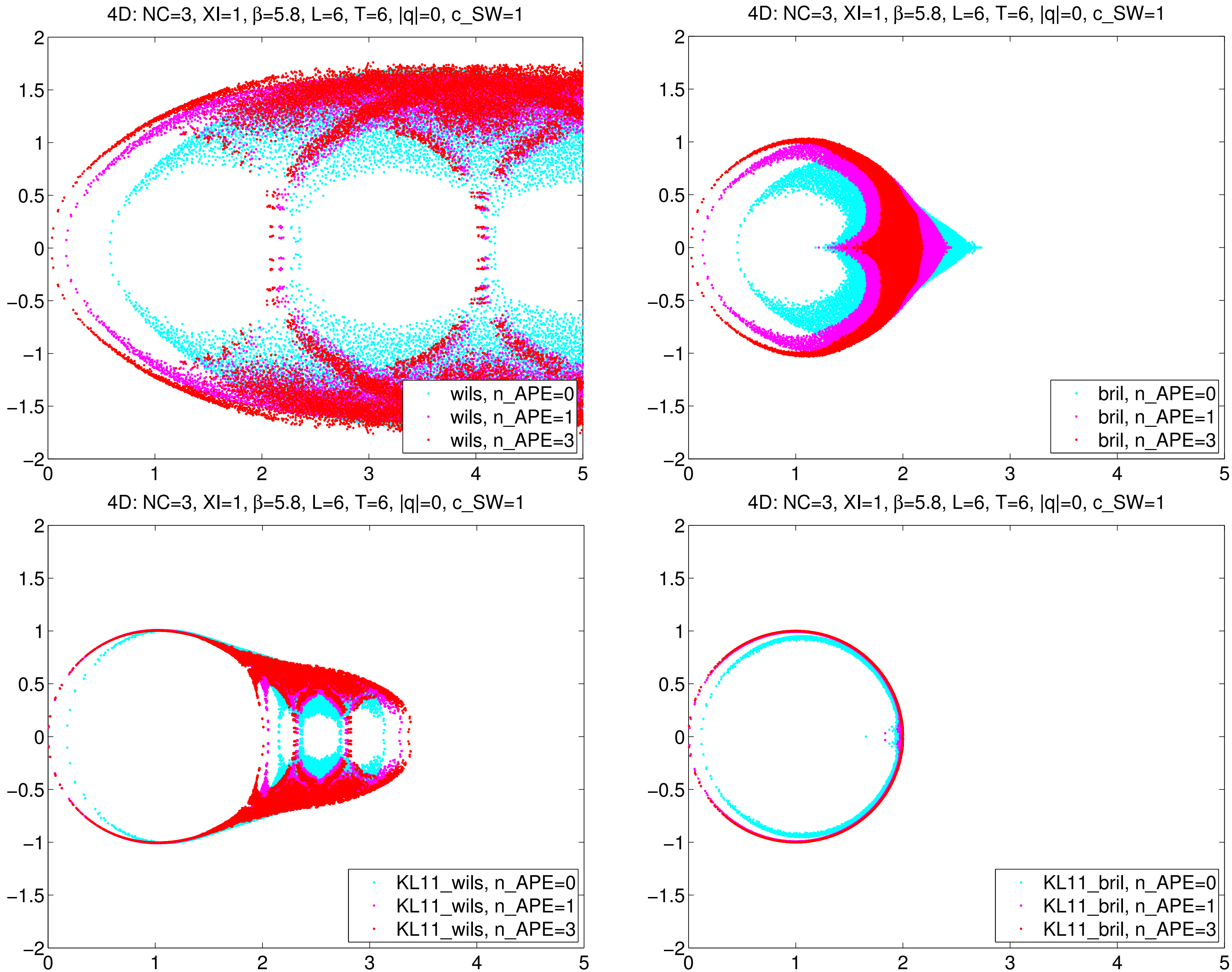
Kenney-Laub family of iterates for matrix sign function / inverse square root:

$$f_{11}(x) = x\frac{3+x^2}{1+3x^2}\,, \quad f_{22}(x) = x\frac{5+10x^2+1x^4}{1+10x^2+5x^4}\,, \quad \ldots\,, \quad f_{11}(f_{11}(x)) = f_{44}(x)$$



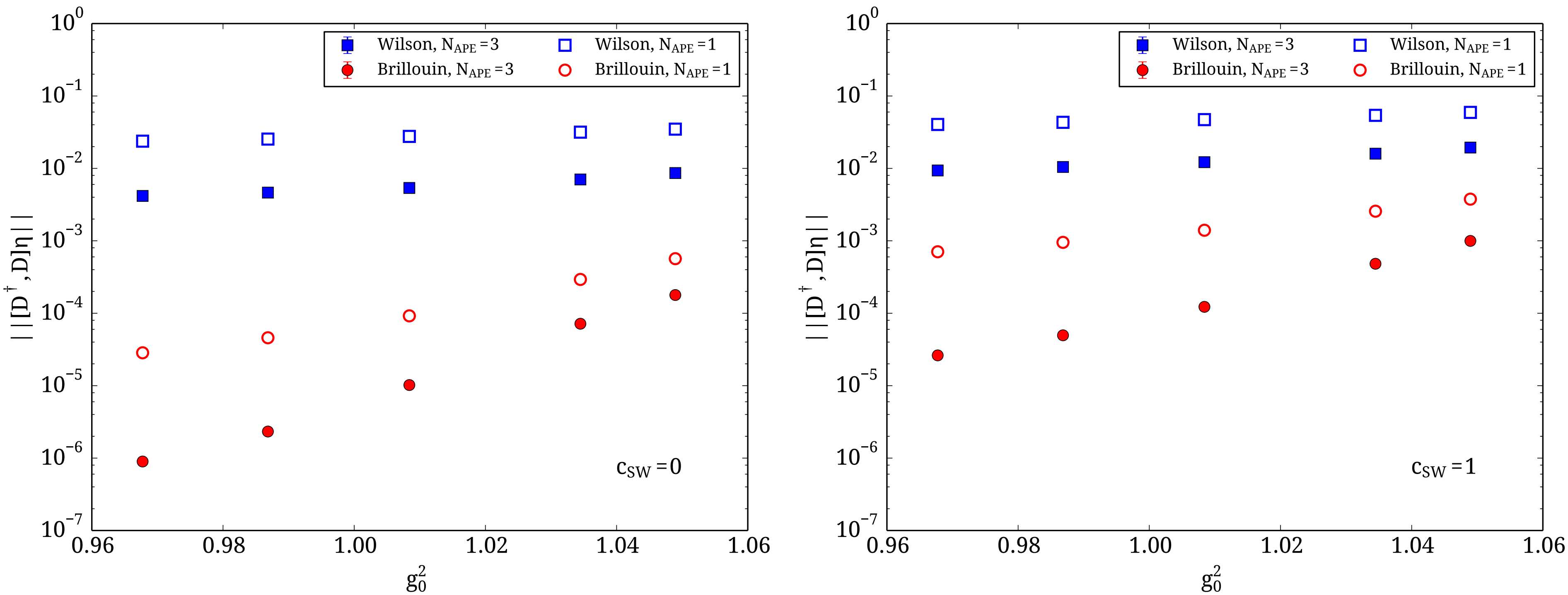
Special features: no wiggles, no eigenvalue info required (but still useful for speed-up)

Other options (Zolotarev polynomial/rational, ...) possible and arguably faster.


4D: NC=3, XI=1, β=5.8, L=6, T=6, |q|=0, c_SW=1
wils, n_APE=0
wils, n_APE=1
wils, n_APE=3
4D: NC=3, XI=1, β=5.8, L=6, T=6, |q|=0, c_SW=1
bril, n_APE=0
bril, n_APE=1
bril, n_APE=3
4D: NC=3, XI=1, β=5.8, L=6, T=6, |q|=0, c_SW=1
KL11_wils, n_APE=0
KL11_wils, n_APE=1
KL11_wils, n_APE=3
4D: NC=3, XI=1, β=5.8, L=6, T=6, |q|=0, c_SW=1
KL11_bril, n_APE=0
KL11_bril, n_APE=1
KL11_bril, n_APE=3

Brillouin overlap: Spectroscopy tests

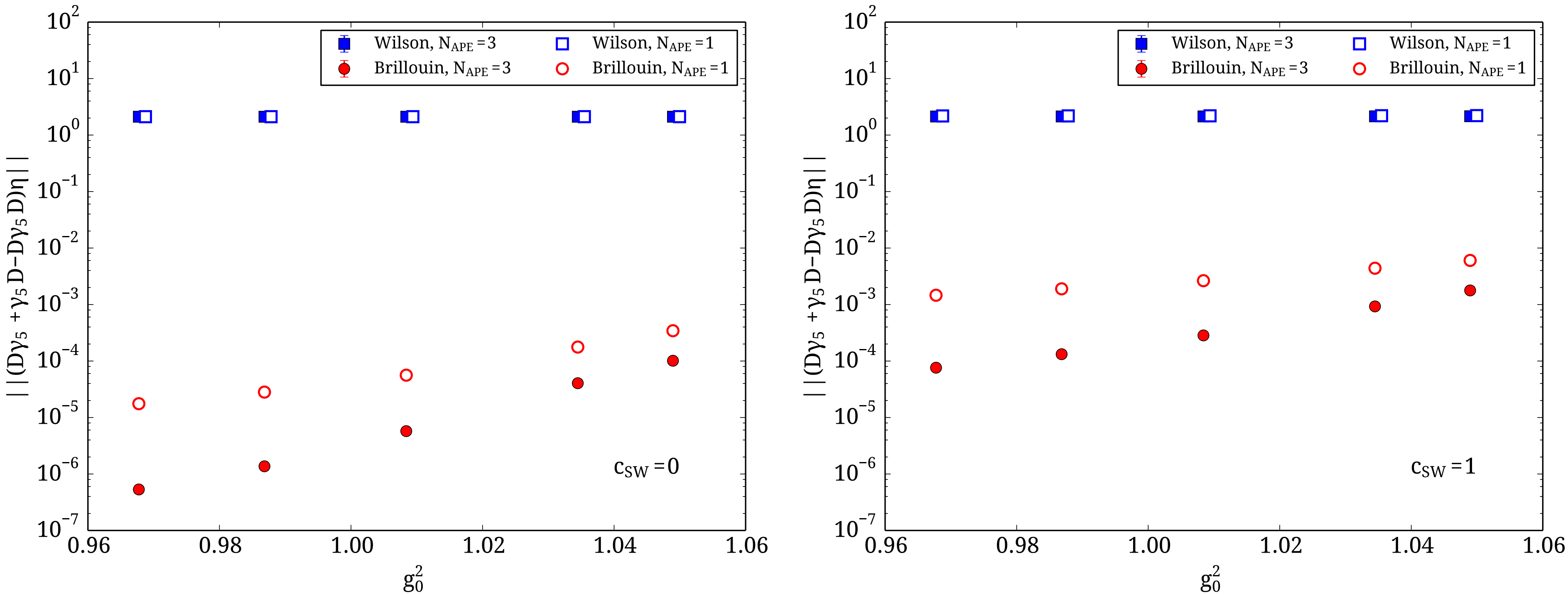
- **Normality**



$\Longrightarrow$ B-overlap closer to normality than W-overlap, and advantage grows for $a \to 0$

Note: here “overlap” denotes the fixed-order (low-grade) approximation KL11

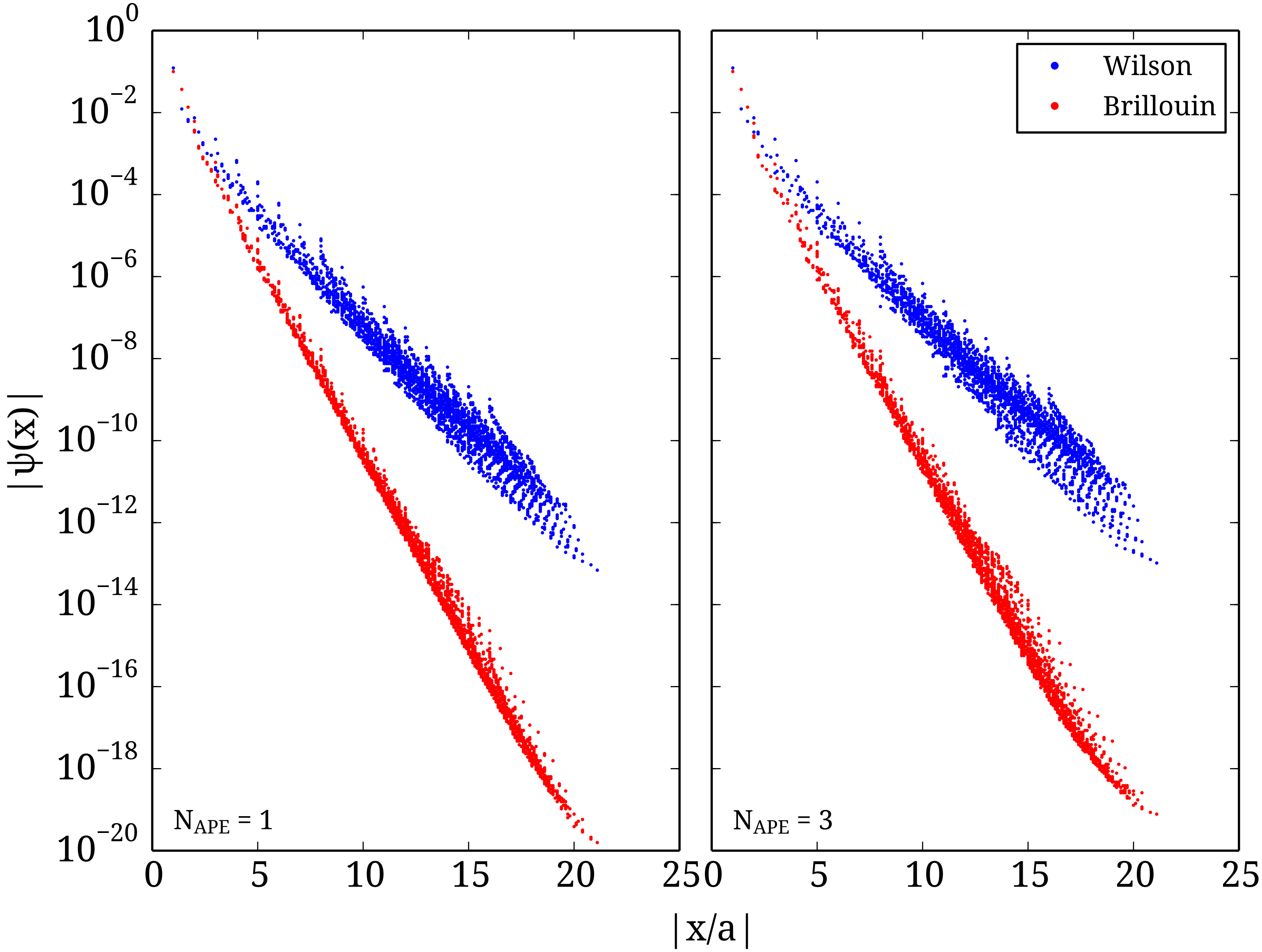
• GW defect



$\Longrightarrow$ B-overlap much more chiral than W-overlap, and advantage grows for $a \to 0$

Note: here "overlap" denotes the fixed-order (low-grade) approximation KL11

- **Operator locality**

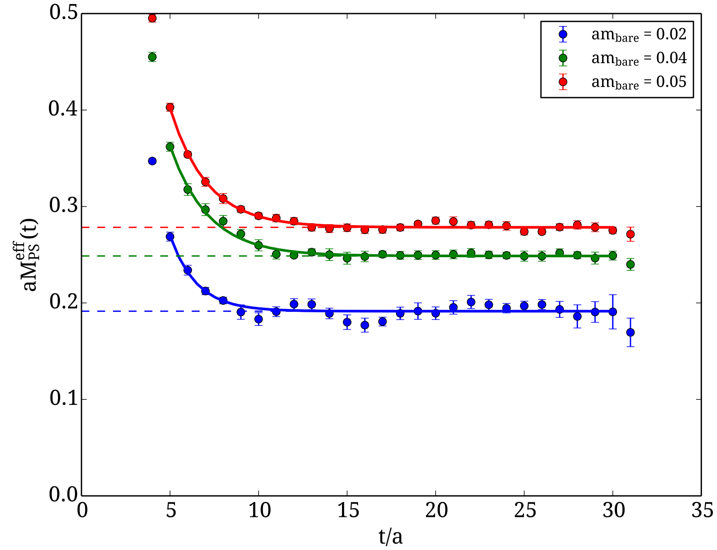


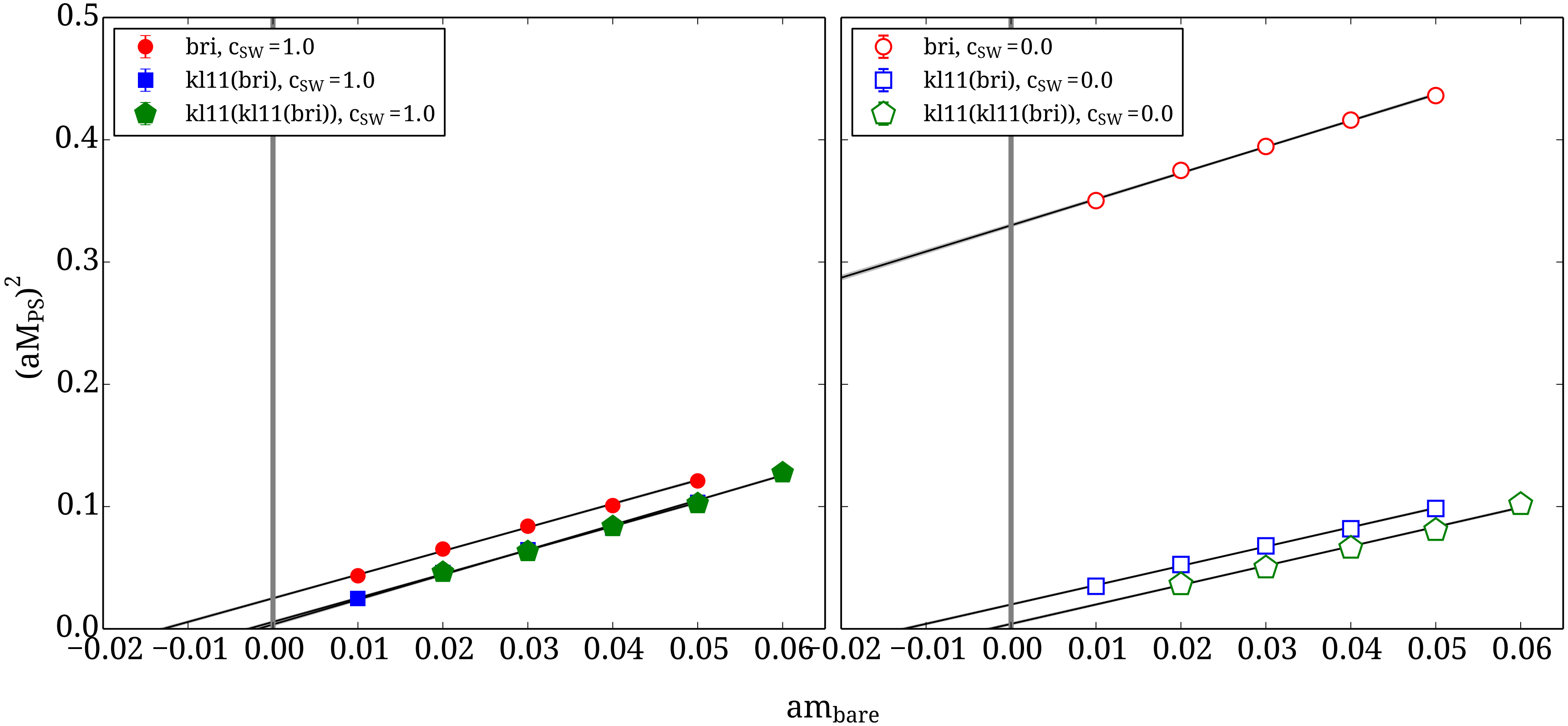
$\Longrightarrow$ B-overlap much better localized than W-overlap, and better rotational symmetry

- **Meson spectroscopy**

D_{kern} (red), D_{KL11} (blue), D_{KL44} (green)
$c_{\text{SW}} = 1$ (left) and $c_{\text{SW}} = 0$ (right)
3 APE smearings throughout

$40^3 \times 64$ lattices, various bare quark masses
Inversion with BiCGstab





Summary

- Brillouin action shows better [more continuum-like] dispersion relation than Wilson action for large momenta but *not* for large quark masses.

- Overlap recipe improves [shown for two kernels, conjectured for any kernel] dispersion relation for large quark masses but *not* for large momenta.

- Brillouin-Overlap/DomainWall yields down-to-earth approximation to "perfect action" concept by Peter Hasenfratz and collaborators, which is universally applicable for small/large quark masses, as well as small/large momenta.

- Implementation details:
  Cost of matrix-vector with Brillouin is 10-20 times higher than with Wilson kernel.
  About a factor 2-4 comes back from reduced iteration count for non-chiral action.
  About a factor 2-4 comes back from reduced iteration count in $\gamma_5 \mathrm{sign}(\gamma_5 D_{-\rho})$.

o *Cheap it is not, but the convenience may be worth it !!!*

Dürr and Koutsou arXiv:1012.3615/1108.1650/1208.6270/forthcoming
Cho et al arXiv:1504.01630, Gattringer et al, Bietenholz et al, Hasenfratz et al