The anomalous magnetic moment of the muon

Anian Altherr, Paola Tavella, Roman Gruber
Institute of Theoretical Physics
ETH Zürich

January 19, 2022

Overview

1. Lattice QCD
2. Introduction to g-2 & Background
3. Hadronic contributions

Non-abelian gauge theories

Yang-Mills Lagrangian (4D, with P and T invariance):

$$\mathcal{L} = -\frac{1}{4} F^{a}_{\mu\nu} F^{\mu\nu,a} + \bar{\psi} \left(i \not{D} - m \right) \psi$$

$$\begin{aligned} F^{a}_{\mu\nu} &= \partial_\mu A^{a}_{\nu}(x) - \partial_\nu A^{a}_{\mu} + g f^{abc} A^{b}_{\mu}(x) A^{c}_{\nu}(x) \\ D_\mu &= \partial_\mu + i g A^{a}_{\mu}(x) T^{a}, \end{aligned}$$

with action $\mathcal{S} = \int d^4x \mathcal{L}$. In path integral formalism, we have the Minkowski expectation value

$$\begin{aligned} \langle \mathcal{O}[\psi, \bar{\psi}, A] \rangle &= \lim_{t \to \infty(1 - i\epsilon)} \frac{1}{Z} \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A \mathcal{O}[\psi, \bar{\psi}, A] e^{iS[\psi, \bar{\psi}, A]} \\ Z &= \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A e^{iS[\psi, \bar{\psi}, A]}. \end{aligned}$$

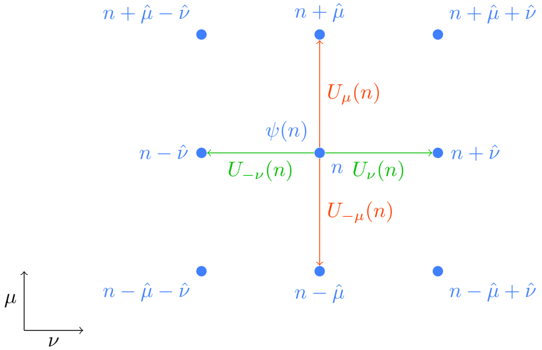
Discretization

4D lattice

$$\Lambda = \{n = (n_0, n_1, n_2, n_3) \mid n_i \in \{0, 1, \dots, L_i - 1\}, i \in \{0, 1, 2, 3\}\}.$$

$$\begin{aligned} A^b_\mu(x) = A^b_\mu(an) &\longrightarrow aA^b_\mu(n), \\ \psi(x) = \psi(an) &\longrightarrow a^{3/2}\psi(n) \\ U(x, x + \epsilon\hat{\mu}) = U(an, an + a\hat{\mu}) &\longrightarrow U_\mu(n). \end{aligned}$$

- lattice extents $L_i \in \mathbb{N}$
- lattice constant a
- compensator field (link variable) $U(x, y)$ $\left(\frac{\partial U}{\partial x^\mu} = igA^a_\mu(x)T^a\right)$



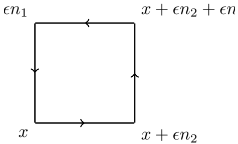
Discretization

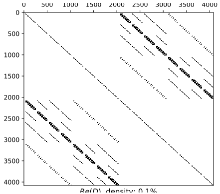
The full discretized Yang-Mills action is

$$\mathcal{S} = \mathcal{S}_G + \mathcal{S}_F,$$

$$\mathcal{S}_G = \frac{1}{g^2} \sum_{n \in \Lambda} \sum_{\mu,\nu} Re\ tr \left[\delta_{\mu\nu} \cdot id - \hat{U}_{\mu\nu}(n) \right],$$

$$\mathcal{S}_F = a^4 \sum_{n \in \Lambda} \bar{\psi}(n) \underbrace{\left[\sum_{\mu} \frac{\gamma^\mu}{2}(D_{+\mu} + D_{-\mu}) + m \right]}_{D} \psi(n)$$





- $D_{\pm\mu}$ are forward- and backward covariant derivatives
- D is the Dirac operator
- Large sparse matrix, dim $d = 12V$, (eg. 64^4 lattice: $d \approx 10^8$, naively 500 petabytes)

Wick rotation

Rotating to the Euclidean action $e^{i\mathcal{S}} \xrightarrow{\text{WR}} e^{-\mathcal{S}_E}$ we obtain the Euclidean expectation value

$$\langle \mathcal{O}[\psi, \bar{\psi}, A] \rangle = \lim_{\tau \to \infty} \frac{1}{Z} \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A \mathcal{O}[\psi, \bar{\psi}, A] e^{-S[\psi, \bar{\psi}, A]}$$

$$Z = \int \mathcal{D}A \det(D) e^{-S_G[A]}.$$

Interpret the exponential as probability density, $P(U) = Z^{-1} e^{-S}$,

$$\langle \mathcal{O}[U] \rangle = \lim_{\tau \to \infty} \frac{1}{Z} \int \mathcal{D}U e^{-S[U]} \mathcal{O}[U] \longrightarrow \sum_U P(U) \mathcal{O}[U],$$

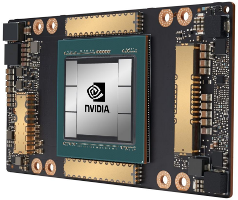
$$\langle \mathcal{O}[U] \rangle \approx \frac{1}{N} \sum_{i=1}^{N} \mathcal{O}[U_i].$$

Lattice QFT on the CPU

- large sparse linear operator (complex dimension $O(10^8)$)
- Markov chain alters the operator
- operator needs to be inverted repeatedly (solve $D\psi = \eta$)
- most used compute intensive kernel: **sparse matrix-vector multiplication** (SpMV), $D\psi$
- sparse linear algebra is **memory bound** (speed limited by memory bandwidth, not compute power)
- using supercomputers ($O(100)$ nodes with $O(100)$ cores) $\rightarrow$ high parallelizability
- large problem is decomposed into many smaller problems

Lattice QFT on the GPU

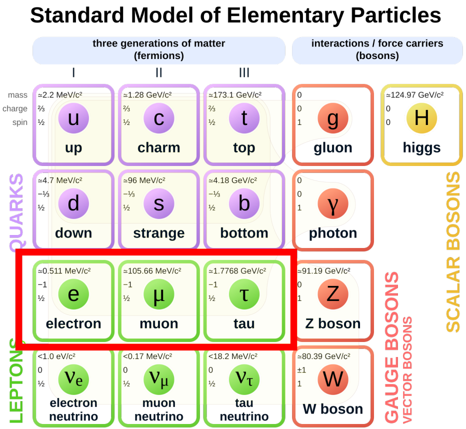
- modern supercomputers tend to include **graphical processing units** (GPUs)
- memory throughput is higher on GPUs
  - Piz Daint CPU node; 52.2 GB/s (Cray XC30)
  - Piz Daint GPU node; 250 GB/s (NVIDIA Tesla K20X)
  - Alps CPU node; ??? GB/s (HPE Cray EX)
  - Alps GPU node; 2039 GB/s (NVIDIA Tesla A100 via NVlink)
- GPU include $O(1000)$ of cores (called steaming multiprocessors (NVIDIA))
- enable GPU processing in openQ*D (based on openQCD)
- enable GPU techniques (tensor-core, reduced precision, ...)



PhD Project

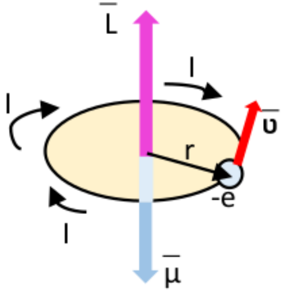
- My PhD is divided into 2 parts
- 1. part: funded via PASC project
  - computer science
  - high performance computing
  - actual implementations
- 2. part: purely theoretic, physics related, g-2, see Paola and Anians presentation

Standard model



Magnetic dipole moment

- particle with mass m,
- electric charge Qe,
- orbiting with angular momentum $\vec{L}$,
- Bohr magneton $\mu_B = \frac{e\hbar}{2mc}$.



Magnetic moment

$$\vec{\mu}_m = \frac{Q\mu_B}{\hbar}\vec{L}.$$

Magnetic dipole moment

- particle with mass m,
- electric charge Qe,
- intrinsic spin $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$,
- Bohr magneton $\mu_B = \frac{e\hbar}{2mc}$.
- gyromagnetic ratio g.

Magnetic moment

$$\vec{\mu}_m = g\,\frac{Q\mu_B}{\hbar}\vec{S}.$$

Anomalous magnetic moment

- Kronig, Goudsmit, Uhlenbeck (1925): $g = 1$,
- Dirac (1928): $g = 2$,
- Experiments: $g \approx 2 + 0.002$.

Anomalous magnetic moment

for lepton ℓ with gyromagnetic ratio g_ℓ:

$$a_\ell := \frac{g_\ell - 2}{2}.$$

Anomalous magnetic moments

	e	μ	τ
mass	0.51 $\mathrm{MeV/c^2}$	110 $\mathrm{MeV/c^2}$	1800 $\mathrm{MeV/c^2}$
lifetime	stable	$2.2 \cdot 10^{-6}$ s	$2.9 \cdot 10^{-13}$ s

[1] Hanneke et al. (2008, 2011), CODATA recommended values of the fundamental physical constants: 2018

Anomalous magnetic moments

	e	μ	τ
mass	$0.51\ \mathrm{MeV/c^2}$	$110\ \mathrm{MeV/c^2}$	$1800\ \mathrm{MeV/c^2}$
lifetime	stable	$2.2 \cdot 10^{-6}\ \mathrm{s}$	$2.9 \cdot 10^{-13}\ \mathrm{s}$

- a_τ hard to determine due to short lifetime.

[1] Hanneke et al. (2008, 2011), CODATA recommended values of the fundamental physical constants: 2018

Anomalous magnetic moments

	e	μ	τ
mass	$0.51\ \mathrm{MeV/c^2}$	$110\ \mathrm{MeV/c^2}$	$1800\ \mathrm{MeV/c^2}$
lifetime	stable	$2.2 \cdot 10^{-6}\ \mathrm{s}$	$2.9 \cdot 10^{-13}\ \mathrm{s}$

- a_τ hard to determine due to short lifetime.
- a_e determined up to less than 1 part per billion both in theory and experiment.[1]

[1] Hanneke et al. (2008, 2011), CODATA recommended values of the fundamental physical constants: 2018

Anomalous magnetic moments

	e	μ	τ
mass	$0.51\ \mathrm{MeV/c^2}$	$110\ \mathrm{MeV/c^2}$	$1800\ \mathrm{MeV/c^2}$
lifetime	stable	$2.2 \cdot 10^{-6}\ \mathrm{s}$	$2.9 \cdot 10^{-13}\ \mathrm{s}$

- a_τ hard to determine due to short lifetime.
- a_e determined up to less than 1 part per billion both in theory and experiment.[1]
- Sensitivity δa_ℓ to new physics (Λ ultraviolet cut-off for new physics):

$$\frac{\delta a_\ell}{a_\ell} \sim \frac{m_\ell^2}{\Lambda^2}$$

[1] Hanneke et al. (2008, 2011), CODATA recommended values of the fundamental physical constants: 2018

Anomalous magnetic moments

	e	μ	τ
mass	$0.51\ \mathrm{MeV/c^2}$	$110\ \mathrm{MeV/c^2}$	$1800\ \mathrm{MeV/c^2}$
lifetime	stable	$2.2 \cdot 10^{-6}\ \mathrm{s}$	$2.9 \cdot 10^{-13}\ \mathrm{s}$

- a_τ hard to determine due to short lifetime.
- a_e determined up to less than 1 part per billion both in theory and experiment.[1]
- Sensitivity δa_ℓ to new physics (Λ ultraviolet cut-off for new physics):

$$\frac{\delta a_\ell}{a_\ell} \sim \frac{m_\ell^2}{\Lambda^2}$$

$\rightarrow$ Muon most suitable lepton for discovering new physics.

[1] Hanneke et al. (2008, 2011), CODATA recommended values of the fundamental physical constants: 2018

Contributions

$$a_\mu = a_\mu^{\text{QED}} + a_\mu^{\text{EW}} + a_\mu^{\text{Had}}.$$

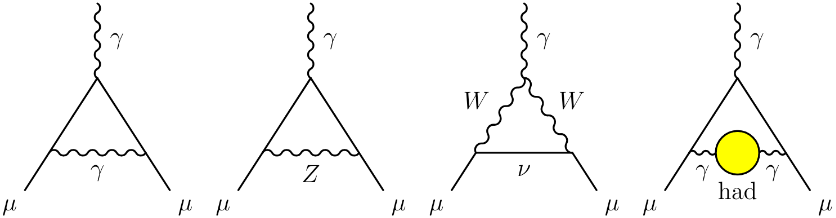


Figure: From left to right: first order QED, lowest-order weak, lowest-order hadronic.

Contributions

Data by Particle Data Group[2]

$$\begin{aligned}
a_\mu^{\text{QED}}[\text{5-loop}] &= 116584718.93(0.10) \cdot 10^{-11}, \\
a_\mu^{\text{EW}}[\text{2-loop}] &= 153.6(1.0) \cdot 10^{-11}, \\
a_\mu^{\text{Had}}[LO] &= 6931(40) \cdot 10^{-11}, \\
a_\mu^{\text{Had}}[N(N)LO] &= 6(18) \cdot 10^{-11}, \\
a_\mu^{\text{SM}} &= 116591810(1)^{\text{EW}}(40)^{\text{Had,LO}}(18)^{\text{Had,N(N)LO}} \cdot 10^{-11}.
\end{aligned}$$

Discrepancy between theory and experiment of 4.2σ:

$$\Delta a_\mu = a_\mu^{\text{exp}} - a_\mu^{\text{SM}} = 251(41)^{\text{exp}}(43)^{\text{SM}} \cdot 10^{-11} \tag{1}$$

[2] Particle Data Group, Review of Particle Physics (2020)

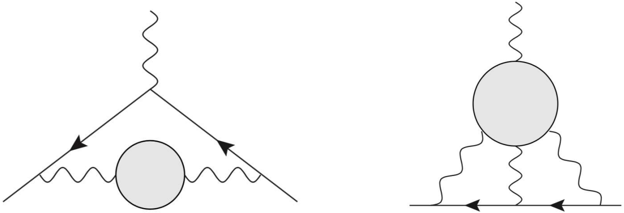
Machine learning and lattice QCD

- Goal: Approximate density $P(U) = e^{-S[\psi,\overline{\psi},U]}/Z$.
- currently: Markov chain Monte Carlo (MCMC) suffers from critical slowing down: Autocorrelation time diverges for large lattice sizes.
- flow-based MCMC overcomes this issue for ϕ^4 theory. [3]

[3] arXiv:1904.12072 (2019)

Hadronic contributions

The noisest contributions are the hadronic ones: Hadronic vacuum polarization (HVP) and hadronic light-by-light (Hlbl)



- at the energy of $O(m_\mu)$ the QCD is non-perturbative
- at lowest order HVP and Hlbl contributions are $O(\alpha^2)$ and $O(\alpha^3)$
- different approaches to estimate these contributions.

The state of the art

Current SM results:

$$a_\mu^{HVP}(\text{e+e, up to NNLO}) =$$
$$= 6845(40) \cdot 10^{-11}$$

$\Longrightarrow$ Lattice results close to the experimental value, but errors are still too large

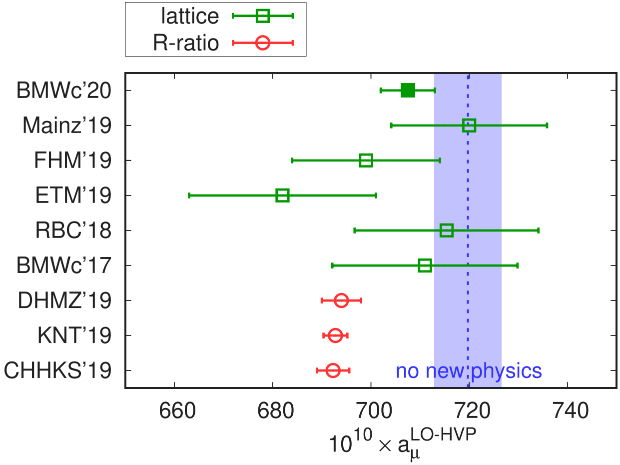
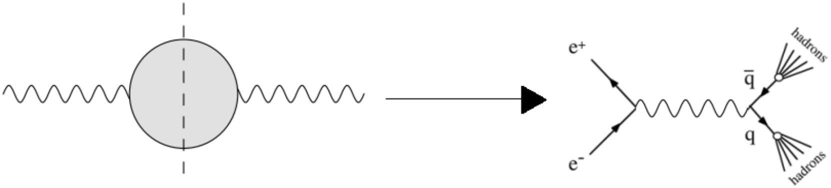


Figure from arXiv:2002.12347 (2021)

Data-driven method



- dispersion relation + optical theorem:

$$\mathrm{Im}\Pi(q^2) \propto \sigma_{e^+e^-\to\text{hadrons}} = R(q^2)\sigma_{e^+e^-\to\mu^+\mu^-} \implies a_\mu^{HVP} = \frac{\alpha^2}{3\pi^2}\int_{M_\pi^2}^{\infty}\frac{ds}{s}K(s)R(s)$$

- use experimental data (BaBar, KLOE, etc.) as input
- achieved precision $\sim 0.6\%$

Lattice approach

- key observable: EM vector-vector correlator (contributions from u, d, s, c, b quarks)

$$\langle \mathcal{O}_{\mu\nu} \rangle = \langle J_\mu(x) J_\nu(0) \rangle$$

$$J_\mu(x) = \sum_f q_f \overline{\psi}^f(x) \gamma_\mu \psi^f(x)$$

- example: by considering 2 flavours and using the Wick contractions:

$$\langle J_k(x) J_k(0) \rangle = \sum_{f,f'} q_f q_{f'} \, \mathrm{tr}\left[\gamma_k D_f^{-1}(x|x)\right] \cdot \mathrm{tr}\left[\gamma_k D_{f'}^{-1}(0|0)\right] - \sum_f q_f^2 \, \mathrm{tr}\left[\gamma_k D_f^{-1}(x|0) \gamma_k D_f^{-1}(0|x)\right]$$

disconnected ~ 2%

connected

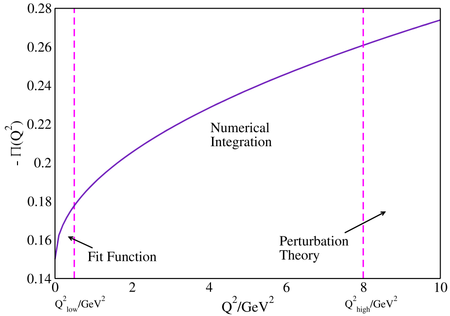
Hadronic vacuum polarization

- introduce the polarization tensor and its decomposition

$$\Pi_{\mu\nu}(Q) = \int d^4x e^{iQ\cdot x}\langle J_\mu(x)J_\nu(0)\rangle = \left(Q_\mu Q_\nu - \delta_{\mu\nu}Q^2\right)\Pi(Q^2)$$

- perform the integration over momenta with a known kernel function[4] $K(Q^2, m_\mu^2)$.

$$a_\mu^{HVP} = 4\alpha^2 \int_0^\infty dQ^2 K(Q^2; m_\mu^2) \cdot \left(\Pi(Q^2) - \Pi(0)\right) = I_1 + I_2 + I_3$$



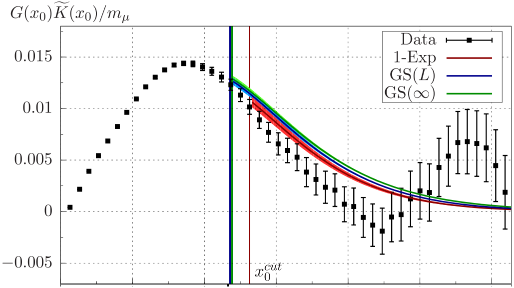
4 arXiv:2006.04822 (2020)

Time-momentum representation (TMR)

- By inverting the order of the spatial Fourier transform and the integration over momenta:

$$G(x_0) = -\frac{1}{3} \sum_{i=1,2,3} \int d^3x \langle J_i(x) J_i(0) \rangle$$

$$a_\mu^{HVP} = \left(\frac{\alpha}{\pi}\right)^2 \int_0^\infty \mathrm{d}t G(t) \tilde{K}(t; m_\mu)$$



- large relative statistical error at large x_0 $\implies$ several models to extrapolate data after x_0^{cut}

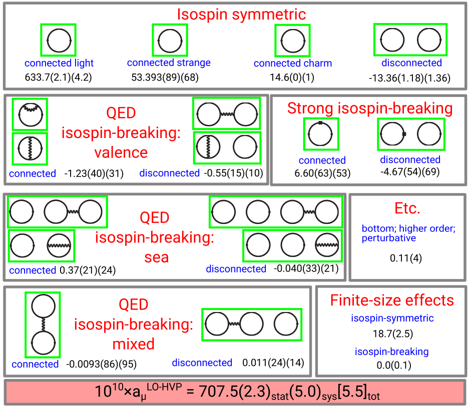
Figure from arXiv:1705.01775 (2017)

Challenges

Many efforts aim to reduce the errors on the theoretical side by focusing on:

- precise determination of the lattice scale a;
- noise reduction technique for statistical errors (smearing, multilevel...);
- finite size effects and continuum limit;
- isospin breaking effects due to QED (α_{EM}) and strong interactions ($m_u - m_d$)

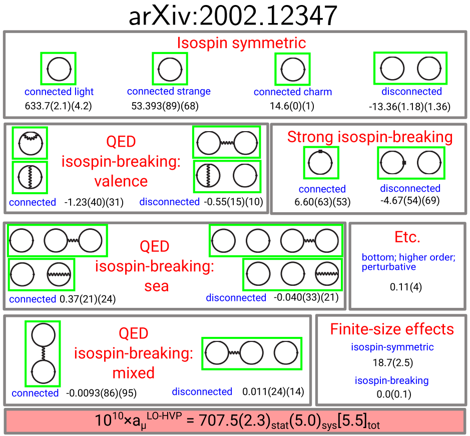
arXiv:2002.12347



Challenges

Many efforts aim to reduce the errors on the theoretical side by focusing on:

- precise determination of the lattice scale a;
- noise reduction technique for statistical errors (smearing, multilevel...);
- finite size effects and continuum limit;
- isospin breaking effects due to QED (α_{EM}) and strong interactions ($m_u - m_d$)



Isospin breaking

1) QCD simulations + perturbative series in α_{EM} and $(m_u - m_d)/(m_u + m_d)$
2) first-principle method: QCD + QED simulations (C^* boundary conditions)

C* boundary conditions

C* bc allows for electrically-charged states and propagation of charged particles

$$A_\mu(x + L_1\hat{e}_1) = -A_\mu(x), \qquad U_\mu(x + L_1\hat{e}_1) = U^*_\mu(x),$$

$$\psi_f(x + L_1\hat{e}_1) = C^{-1}\overline{\psi}_f^T(x), \qquad \overline{\psi}_f(x + L_k\hat{e}_k) = -\psi_f^T(x)C.$$

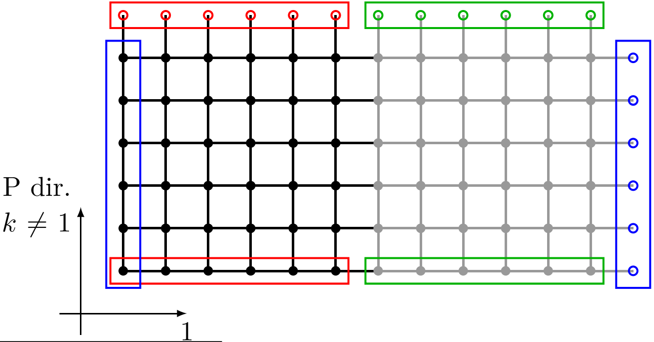


figure from arxiv:1908.11673 (2019)

C* boundary conditions

C* bc allows for electrically-charged states and propagation of charged particles

$$A_\mu(x + L_1\hat{e}_1) = -A_\mu(x), \qquad U_\mu(x + L_1\hat{e}_1) = U^*_\mu(x),$$

$$\psi_f(x + L_1\hat{e}_1) = C^{-1}\overline{\psi}_f^T(x), \qquad \overline{\psi}_f(x + L_k\hat{e}_k) = -\psi_f^T(x)C.$$

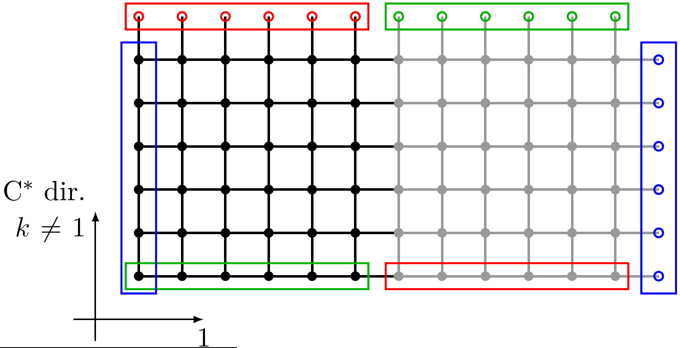


figure from arxiv:1908.11673 (2019)