Massive 1D Dirac Line, Solitons and Reversible Manipulation on the Surface of a Prototype Obstructed Atomic Insulator, Silicon

Z. K. Liu, P. Deng, Y. F. Xu, H. F. Yang, D. Pei, C. Chen, S. M. He, D. F. Liu, S. K. Mo, T. K. Kim, C. Cacho, H. Yao, Z. D. Song, X. Chen, Z. Wang, B. H. Yan, L. X. Yang, A. Bernevig, and Y. L. Chen

Part I **Additional information on the real space invariant analysis of Silicon**

1.1: Obstructed atomic insulator phase in silicon

1.2: Existence of solitons at the interface of NBC and PBC atomic chains

Part II **Details of the 1D Massive Dirac fermion model for surface atomic chains**

2.1: 1D Massive Dirac fermions model

2.2: 0D soliton solutions at the interface between NBC and PBC atomic chains

2.3: A brief comparison between Si (111) surface atomic chains and polyacetylene

Part III **Additional information of *ab initio* calculations on the OSS**

Part IV **Additional discussion on chain domains with different orientations**

Part V **More band dispersions from NBC and PBC domains**

Part VI **Addition information on STM/STS measurements and domain switching**

6.1: More details of interface solitons

6.2: Discussion on the switching of NBC/PBC domains

6.3. Determine ('read') NBC/PBC domains using LDOS information.

Part I

Additional information on the real space invariant analysis of Silicon

1.1: Obstructed atomic insulator phase in silicon

The high-symmetry crystaline silicon with space group $Fd\bar{3}m$ is a topologically trivial insulator. The two Si atoms sit at Wyckoff position 2a(1/8,1/8,1/8). The single-valued real space invariant (RSI)[1, 2] at the Wyckoff position 4d(1/2,1/2,1/2) is calculated as $\delta_6(d) = 1$ (which is defined in Table S1 as: $\delta_6(\mathrm{d}) = 3\mathrm{N}(\Gamma_2^+) - \mathrm{N}(\Gamma_3^-) - \mathrm{N}(\Gamma_4^-) + 2\mathrm{N}(\Gamma_5^-) - \mathrm{N}(\mathrm{L}_1^-) - 2\mathrm{N}(\mathrm{L}_2^+) + \mathrm{N}(\mathrm{L}_3^+)$, where $N(R)$ represents the multiplicity of the irreducible representation R below the Fermi level.

The nonzero RSI $\delta_6(d) = 1$ indicates an obstructed atomic insulator where the Wannier

Single-valued RSI global indices in the space group $Fd\bar{3}m$ (N. 227)

Index	Wyckoff positions	Irreps and multiplicity																					
		Γ_1^+	Γ_1^-	Γ_2^+	Γ_2^-	Γ_3^+	Γ_3^-	Γ_4^+	Γ_4^-	Γ_5^+	Γ_5^-	L_1^+	L_1^-	L_2^+	L_2^-	L_3^+	L_3^-	W_1	W_2	X_1	X_2	X_3	X_4
δ_1	a	0	0	-1	0	0	1	1	0	0	0	1	1	0	0	0	0	0	-1	0	0	0	0
δ_2	b	0	0	1	0	0	-1	0	-1	-1	1	0	-1	-1	0	0	0	0	1	0	0	0	0
δ_3	c	0	0	1	0	0	0	0	0	0	1	0	-1	0	0	0	0	0	0	0	0	0	0
δ_4	c	0	0	1	0	0	1	1	0	1	1	0	-1	0	0	-1	0	0	0	0	0	0	0
δ_5	d	0	0	1	0	0	0	0	0	0	1	0	0	-1	0	0	0	0	0	0	0	0	0
δ_6	d	0	0	3	0	0	-1	0	-1	0	2	0	-1	-2	0	1	0	0	0	0	0	0	0
δ_7	e	0	0	0	0	0	1	1	0	1	0	-1	-1	0	0	0	0	0	0	0	0	0	0
δ_8	f	0	0	1	0	0	-1	0	1	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0
δ_9	g	0	0	-2	0	0	0	-1	1	0	-2	-1	1	2	0	0	0	0	0	0	0	0	0
δ_{10}	h	0	0	-2	0	0	0	0	1	-1	-2	0	1	1	0	0	0	0	0	0	0	0	0

Table S1 Single-valued RSIs (without SOC) defined at all the Wyckoff positions of SG 227.

charge centers (WCCs) of the occupied bands mismatch with the atomic sites. Given this nonzero RSI any cleavage plane cutting the Wyckoff position 4d would give rise to the gapless surface states in the bulk gap. In Figure. S1b, we calculate the surface states of the silicon (111) surface.

By counting the filling number, the obstructed surface states (labelled by the red lines in Figure S1b) on the Fermi level is half filled (filling anomaly). As the RSI is determined by the occupied bulk states and protected by inversion symmetry, weak perturbations (without changing the number of valence electrons) on the surface cannot remove the surface states.

As we consider a 2×1 supercell on the (111) surface (Figure. S1c), the cell expansion (without surface reconstruction) does not change the obstructed atomic insulator phase of the bulk states (because the positions of Si atoms have no inversion symmetry and the Wannier charges are always localized at the inversion invariant center) but induces band folding on the (111) surface. Such band folding results in the band crossing at the boundary of the folded BZ, K', J and J' points, as shown in Figure. S1d.

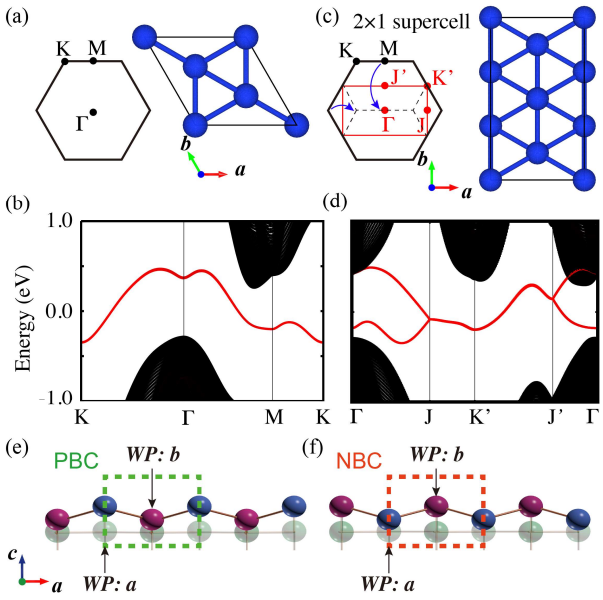


Figure S1 | (a) The crystal structure of Si (111) surface and its projected surface BZ. (b) Calculated bandstructure of the Si (111) surface. The bands marked in red are the obstructed surface states. By counting the filling number of electrons, the obstructed surface state is half-filled. (c) The crystal structure of Si (111) 2x1 supercell and the corresponding surface BZ (red) which is obtained by folding the primitive BZ (black). (d) Calculated bandstructure of the Si (111) 2x1 supercell. The bands marked in red are the obstructed surface states. Band folding in the supercell induces the surface band crossings at K', J and J' points and the band crossings will be gapped by the 2x1 surface reconstruction. (e-f) Schematic of the two types of buckled Si chains due to the 2x1 reconstruction on the Si (111) surface with the maximal Wyckoff positions (WPs) labelled. PBC: positively buckled chain. NBC: negatively buckled chain.

1. 2: Existence of solitons at the interface of NBC and PBC atomic chains

Due to the surface reconstruction, the 2D plane groups of both the NBC and PBC regions on the (111) surface are subducted from *P3m1* to *P1m1*. Since all of the irreducible representations at the high-symmetry k points of *P1m1* are one dimensional (without SOC), the crossing points at K', J and J' are not symmetry protected and can be gapped by the reconstructions, i.e. the positive and negative buckling occuring on the (001) plane of the lattice structure in Fig. S1c, as shown in Fig. 1d in the main text.

In the plane group *P1m1*, there are two distinct maximal Wyckoff positions with point group *m*: the Wyckoff position (WP) *a* at (0, y) and *b* at (½, y), which are defined on the line *x*=0 and *x*=a/2 (a is the lattice constant of the reconstructed surface along *x* direction), respectively. On the top (111) surface, two nonequivalent Si atoms are located at the WP *a* and *b*, respectively. From the slab calculations in Figure S1(d-e), the band structures of both PBC and NBC are gapped at all of the high symmetry points, which allows the definition of RSI at WP *a* and WP *b*. The single-valued RSIs at *a* and *b* of plane group *P1m1* are derived as follows [1],

$$\delta(a) = N(\Gamma_2) - N(K'_1), \quad \delta(b) = N(K'_1) - N(\Gamma_1),$$

BR	N(Γ_1)	N(Γ_2)	N(J_1)	N(J_2)	N(K'_1)	N(K'_2)	N(J'_1)	N(J'_2)	δ(a)	δ(b)
NBC	19	6	19	6	13	12	13	12	-7	-6
PBC	19	6	19	6	12	13	12	13	-6	-7

Table S2. Band representation (BR) of a finit-size slab of Si (including 12 Si atoms) with NBC and PBC on the top (111) surface.

where the mirror eigenvalue of $\Gamma_{1(2)}$ and $K'_{1(2)}$ is +1(-1). By constructing a finite-size slab of Si with PBC(NBC) on the top (111) surface, we have calculated the band representations of its valence bands at the four high-symmetry momenta [2], which are tabulated in Table S2. Using the RSI

formulae defined above, the RSIs of the slabs with NBC and PBC are obtained as {δ(a) = -7, δ(b) = -6} and {δ(a) = -6,δ(b) = -7}, respectively. As the bulk states of the two types of slabs are exactly the same, the differences of RSIs indicate the difference of band topologies on the surface. It means that the valance bands on the surface of PBC and NBC are in the different atomic-limit phases, which cannot adiabatically evolve to each other without closing the bandgap.

From the first principle calculation, energy difference between NBC and PBC is about 3.5 meV per unit cell (less than 1 meV per Si atom on the reconstructed surface) with PBC energetically favorable. Indeed, PBC and NBC are not symmetry related and hence the total energy of them can be different. While in our experiments, as shown in Fig. 3b and Fig. 4 in the main text, NBC and PBC domains can both exist on the reconstructed surface of Si (111). The coexistence of them is owing to the small difference of total energy and a large energy barrier to switch them. Therefore, the domain wall on the boundary between PBC and NBC is relatively stable and the difference of RSIs of the two domains guarantees gapless states (soliton in 1D) on the domain wall due to the closing of the bandgap.

Part II

Details of the 1D Massive Dirac fermion model for surface atomic chains

2.1: 1D Massive Dirac fermions model

The 2×1 reconstruction results in the formation of NBC and PBC chains on the Si(111) surface (see Fig. 1d, Fig. 2a and Fig. 4a of the main text). The main features of the band structure along these chains can be captured by a simple tight-binding model. Consider a unit cell that includes two lattice sites as A and B with a p_z orbital at each site (see Fig. 1d) and omit the weak interchain hopping (given the weak dispersion along the interchain direction, see Fig. 2b in the main text), we get a simplified 1D lattice Hamiltonian:

$$H = -t\sum_i\left[c_{i,A}^{\dagger}\left(c_{i,B}+c_{i-1,B}\right)+h.c.\right]+m\sum_i\left(c_{i,A}^{\dagger}c_{i,A}-c_{i,B}^{\dagger}c_{i,B}\right) \qquad (1)$$

This takes a special form of the Rice-Mele model. For simplicity we have ignored the spin-orbit coupling and omitted the spin index. The k-space form of Eq. (1) is $H = \sum_k c_k^\dagger h(k) c_k$, in which $c_k = (c_{k,A}, c_{k,B})^T$, and

$$h(k) = -t[1 + \cos(ka)]\sigma_x - t \cdot sin(ka)\sigma_y + m\sigma_Z \qquad (2)$$

where *a* is the lattice constant of the one-dimensional chain, and σ_i $(i = x, y, z)$ are Pauli matrices, $\sigma_Z = 1$ (-1) referring to the *A* (*B*) sublattice. In the Hamiltonian of Eq. (2), we have set the hopping parameters $t_{i,i-1}^{A,B} = t_{i,i}^{A,B} = t$, where $t_{i,i-1}^{A,B}$ is the kinetic hopping from A site in unit cell i to B site in unit cell $i-1$. When m=0, h(k) is gapless at $k = \pi/a$ corresponding to the case without surface reconstruction on the (111) surface of Si. Near $k = \pi/a$, the low energy Hamiltonian can be approximated by

$$h(q) = v_F q \sigma_y + m\sigma_Z \qquad (3)$$

where $v_F = at$ and $q = k - \pi/a$. This simple Hamiltonian describes one-dimensional massive Dirac fermions. In our present problem, as confirmed by our *ab initio* calculations, the PBC is described by a positive mass $m = m_1 \approx 0.2$ eV, while the NBC is described by a negative mass $m = m_2 \approx -0.1$ eV.

Similar simplified Hamiltonians arises in the description of polyacetylene and related models. For the polyacetylene, a lattice Hamiltonian is $H = \sum_i (t_1 c_{i,A}^\dagger c_{i,B} + h.c.) + \sum_i (t_2 c_{i+1,A}^\dagger c_{i,B} + h.c.)$, whose low energy manifestation is $h(q) = v_F q \sigma_x - m\sigma_x$, where $m = t_1 - t_2$. Therefore, our system differs from polyacetylene in that the mass term for the latter is $m\sigma_x$ instead of $m\sigma_z$. The $m\sigma_x$ mass describes staggered hopping amplitude, while the $m\sigma_z$ describes staggered on-site energy. In high energy physics, it is well known that there are two types of fermions mass, the parity-preserving mass $\bar{\psi}\psi = \bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L$ and the parity-violating "γ^5 mass" $i\bar{\psi}\gamma^5\psi = i(\bar{\psi}_L\psi_R - \bar{\psi}_R\psi_L)$. In our model, the parity transformation is $\psi \longrightarrow \sigma_x\psi$, which interchanges the A and B sublattices, therefore, $m\sigma_z$ is the parity-violating mass. In polyacetylene, the mass term preserves parity.

A less obvious distinction between Eq. (1) and the polyacetylene model is related to the topological classification of insulators. The polyacetylene is classified as a non-trivial topological insulator of the chiral class in 1D. Because the simultaneous presence of σ_x, σ_y, and σ_z, Eq.(1) lacks the chiral symmetry of polyacetylene, thus it cannot be classified as a nontrivial topological insulator. This difference has subtle effects on the topological zero modes – although we do remark that chiral symmetry, even for polyacetylene, is not a good microscopic symmetry.

2.2: 0D soliton solutions at the interface between NBC and PBC atomic chains

In this part, we study possible zero modes in the presence of topological solitons. First let us study a mass domain wall in the continuum limit, namely, we solve the low energy Hamiltonian

$$h(x) = iv_F\sigma_y\partial_x + m(x)\sigma_Z \qquad (4)$$

in which $m(x)$ is a smooth function of x. A domain wall is realized if $m(x) \longrightarrow m_1$ as $x \longrightarrow \infty$ and $m(x) \longrightarrow m_2$ as $x \longrightarrow -\infty$, with the signs of m_1 and m_2 being the opposite. In our problem, we have $m_1 > 0$ (positive buckling) and $m_2 < 0$ (negative buckling). The normalizable (quasi-)zero mode solution can be obtained as

$$|\psi_0\rangle \propto \binom{1}{-1}\exp\left[-\int_0^x dx' \frac{m(x')}{v_F}\right] \qquad (5)$$

To be more concrete and precise, we proceed to find the (quasi-)zero modes in the lattice model described by Eq. (1). We study a sharp domain wall between positive and negative buckling chains, where two types of domain walls may exist. As illustrated in Fig. 1f, one case exhibits two higher Si atoms, called up kink, in which the mass in the n-th unit cell is taken as $m(n) \longrightarrow m_1$ for $n \geq 1$ (the PBC side), while $m(n) \longrightarrow m_2$ for $n \leq 0$ (the NBC side, see Fig. 1f(i-ii)). The other case has two lower Si atoms, called down kink, also with $m(n) \longrightarrow m_1$ for the PBC side and $m(n) \longrightarrow m_2$ for the NBC side (see Fig. 1f(iii-iv)). We take the down kink as an example. Its energy eigenvalue equations read

$$\big(E + m(n)\big)\psi_{nB} = -t\big(\psi_{nA} + \psi_{n+1,A}\big),$$

$$\big(E - m(n)\big)\psi_{nA} = -t\big(\psi_{nB} + \psi_{n+1,B}\big) \qquad (6)$$

To solve these equations we take the ansatz

$$\psi_{nA} = \psi_A^R \lambda_1^{n-1},\ \psi_{nB} = \psi_B^R \lambda_1^{n-1},\ (\mathrm{n} \geq 1) \qquad (7)$$

$$\psi_{nA} = \psi_A^L \lambda_2^{n},\ \psi_{nB} = \psi_B^L \lambda_2^{n},\ (\mathrm{n} \leq 0) \qquad (8)$$

in which $\lambda_1,\ \lambda_2,\ \psi_A^R,\ \psi_B^R,\ \psi_A^L,\ \psi_B^L$ are unknowns to be found. To ensure that the wavefunction is normalizable, we must have $|\lambda_1| < 1$ and $\lambda_2 > 1$. Inserting this ansatz into Eq. (6), we have

$$(E + m_1)\psi_B^R = -t\psi_A^R(1 + \lambda_1) \quad ,$$

$$(E - m_1)\psi_A^R = -t\psi_B^R\left(1 + \frac{1}{\lambda_1}\right) \qquad (9)$$

which follows from Eq. (6) with $n \geq 2$, and

$$(E + m_2)\psi_B^L = -t\psi_A^L(1 + \lambda_2) \quad ,$$

$$(E - m_2)\psi_A^L = -t\psi_B^L\left(1 + \frac{1}{\lambda_2}\right) \qquad (10)$$

which follows from Eq. (6) with $n \leq -1$, and finally

$$(E - m_1)\psi_A^R = -t(\psi_B^R + \psi_B^L) \quad ,$$

$$(E + m_2)\psi_B^L = -t(\psi_A^R + \psi_A^L) \qquad (11)$$

which are the boundary conditions of domain wall. Solving Eq. (9), (10), and (11), we are able to find the quasi-zero mode energy (to the lowest order of $m_1,\ m_2$):

$$E \approx -\frac{m_1 m_2}{2t}, \qquad (12)$$

and the wave function decay factors

$$\lambda_1 \approx -\left(1 - \frac{m_1}{t}\right), \lambda_2 \approx -\left(1 - \frac{m_2}{t}\right) \qquad (13)$$

The corresponding wave function exhibits the form of a bonding state by two Si atoms of the down kink, as confirmed by our *ab initio* calculation (see Fig. 1f(iv)). We have $t \approx \frac{v_F}{a} \approx 0.8$ eV, therefore, the quasi-zero mode energy is $E \approx 0.013$ eV. We can obtain the quasi-zero mode energy in the up-kink case:

$$E \approx \frac{m_1 m_2}{2t}, \qquad (14)$$

where the wave function displays the form of an anti-bonding state at the up kink (see Fig. 1f(ii)).

This mode does not have exact zero energy, unlike the result of continuum approximation. Nonetheless, it is qualified to be called a "quasi-zero mode" because, in the $|m_i/t| \ll 1$ limit, we have $E \ll |m_i|$, namely, the quasi-zero mode energy is parametrically suppressed compared to the bulk gap $|m_i|$. The condition that $|m_i/t| \ll 1$ is satisfied in our present problem of silicon surface.

2.3: A brief comparison between Si (111) surface atomic chains and polyacetylene

Due to the chiral symmetry, for the polyacetylene model, a mass domain wall accommodates a zero mode with exact zero energy, in both the low energy approximation of Dirac Hamiltonian and the lattice calculation. The existence of exact zero mode is consistent with the topological classification of insulators and the idea of topological protection. In our model, there is no topological protection of zero mode analogous to the case of polyacetylene, instead, the existence of quasi-zero mode is guaranteed by the condition $|m_i/t| \ll 1$. In this limit, the quasi-zero mode energy is at most of the order of m^2/t. In other words, the existence of quasi-zero mode in our model is a consequence of massive Dirac fermions with sign changing gap.

The difference between our model and the polyacetylene can also be appreciated in the open chain case. At an end of an open chain of polyacetylene, there exists an exact zero mode in the regime $m \equiv t_1 - t_2 < 0$. It is unimportant whether m is small compared to t or not. In other words, it does not matter whether the energy bands mimic that of Dirac fermion or not. In contrary, we have checked that the ends of an open chain in our system, with Hamiltonian given in Eq.(1), do not have zero or quasi-zero mode. The approximation of Dirac fermion breaks down at the ends, where the hopping abruptly drops from t to zero. However, the presence of an in-gap state is guaranteed by the different RSIs of the two atomic insulators.

Part III

Additional information of *ab initio* calculations on the OSS

The 2×1 reconstruction forms NBC and PBC chains on the Si (111) surface (see Fig. 1d), which are buckled out-of-plane in opposite ways. Thus the two atoms within each unit cell are not fully equivalent. We first revisit the surface band structure by *ab initio* calculations within the framework of density-functional theory (DFT) and local density approximation (LDA). We employed the Vienna ab initio simulation package with a plane wave basis [3]. The core electrons were represented by the projector-augmented-wave potential[4].

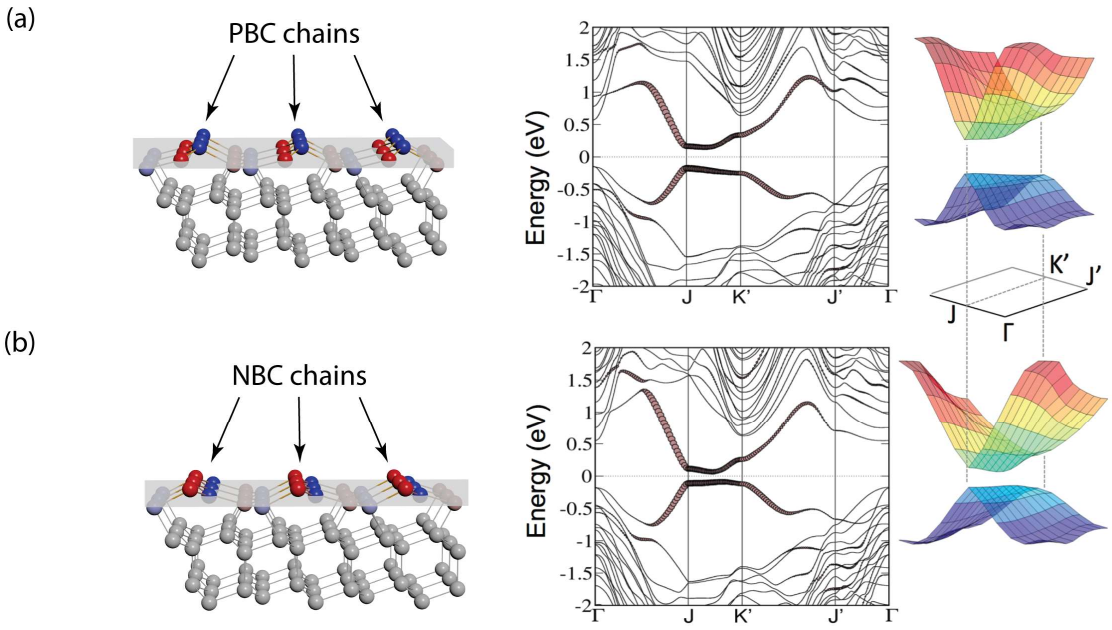
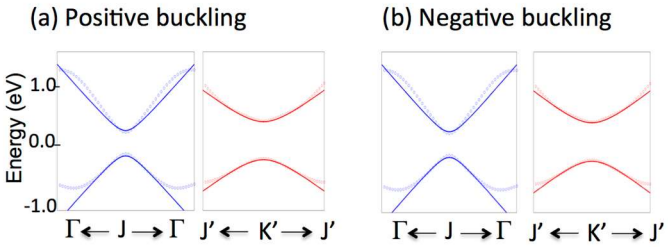


Figure S2 | Band structure of (a) positive and (b) negative buckling surfaces. The size of the filled circles stands for the contribution of Si-p_z orbitals in the atomic chain. The Fermi energy is set to zero. The three-dimensional plot of surface bands are shown in the surface Brillouin zone on right panels, where 1D massive Dirac states can be clearly observed along the chain direction.

The Si(111)-2×1 surface is simulated by a 12-atomic-layer slab (24 Si atoms), where the bottom surface is passivated by hydrogen atoms. Atomic positions of topmost eight atomic layers and hydrogen atoms are fully optimized before the band structure calculations.

The *ab initio* band structures of Si(111)-2×1 surface are shown in Fig. S2. For both positive and negative buckling domains, quasi-1D bands appear inside the bulk energy gap. The dispersion is much stronger along the chain direction than that perpendicular to the chain, consistent with its quasi-1D nature. According to the charge density analysis, the highest occupied surface band at the J point is dominated by the p_z orbital of the higher Si atom while the lowest unoccupied surface band at J is dominated by the p_z orbital of the lower Si atom, for both negative and positive buckling cases. This is consistent with previous calculations [5-7]. Because the surface bands are close to being linear in a wide range of Brillouin zone, we fit the band dispersion to massive Dirac fermions, and show optimal fitting parameters in Fig.S3.



(c)	Positive buckling		Negative buckling	
Dirac point	J	K'	J	K'
v_F (eV*Å)	3.0	1.8	3.3	1.8
m (eV)	0.2	0.3	-0.1	-0.2

Figure S3 | 1D massive Dirac cones exist along the chain direction ($\Gamma - J$ and $J' - K'$). The ab initio bands (empty circles) are fitted according to Eq.4 in Part II (solid lines) near the J and K' points for two surfaces in (c).

Part IV

Additional discussion on chain domains with different orientations

Due to the C_3 rotation symmetry of Si (111) surface, the 2×1 reconstruction can form atomic chains along three equivalent directions, as illustrated in Fig. S4a (note that the red and blue atoms

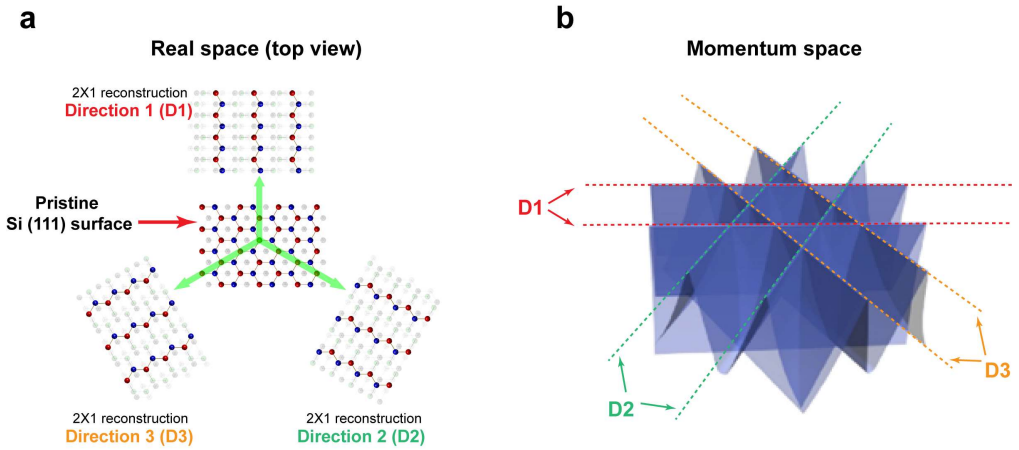


Figure S4 | a, Schematic shows that atomic chains along three equivalent directions can be formed by the 2×1 reconstruction on Si (111) surface. **b,** Schematic shows the superposition of three sets of quasi-1D massive Dirac dispersions corresponding to the three atomic chain directions in **a**. Red dotted lines illustrate the Fermi-surface.

show slightly different out-of-plane position, see Fig. 1d in the main text for more details). Domains of different chain directions can be seen in real space by STM topography mapping as demonstrated in Fig. 2a(ii) of the main text; or in momentum space by ARPES measurements as illustrated in Fig. 2c(i), in which the Kagome shaped Fermi-surface is fomed by the superposition of three sets of quasi-1D massive Dirac dispersions (corresponding to the three chain directions), as shown by the schematic in Fig. S4b.

Part V

More band dispersions from NBC and PBC domains

To discriminate between NBC and PBC domains, we can make use of the difference in their respective band dispersions. As discussed in Part III, their band velocity and gap size both show clear difference (see Fig. S2, S3, due to their slightly different interaction to the bulk atoms), which can be observed in ARPES measurements.

Using the dispersion as an example, in addition to Fig. 3 of the main text, in Fig. S5 we show more ARPES dispersions collected from different sample positions (Fig. S5a) and their extracted band velocities (Fig. S5b). Clearly, the band velocities can be divided into two groups (with the average *hv*=3.3 eV·Å and *hv*=2.8 eV·Å, respectively), corresponding to the dispersions from the

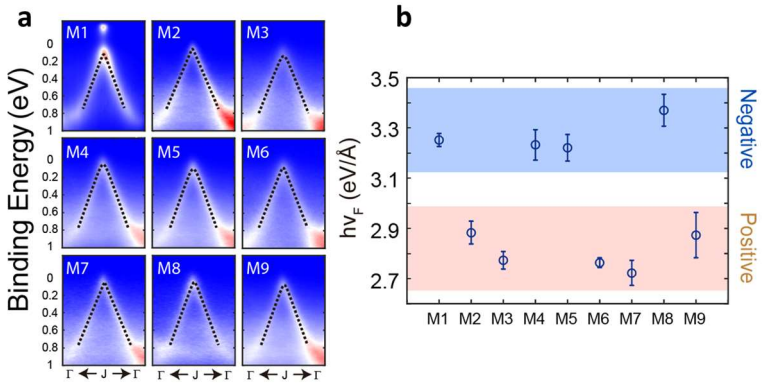


Figure S5 | a, Nine dispersion along $\Gamma - J - \Gamma$ directions measured from different surface positions. Dotted lines shows the fitted dispersions **b,** Extracted band velocity from the dispersions in **a**. Blue and pink shaded regions mark the band velocities from the NBC and PBC domains, respectively.

.

NBC and PBC domain, respectively. Beside ARPES measurements, the difference of NBC and PBC domains can also be seen in STS measurements, as demonstrated in Fig. 4 a(iii,iv) of the main text and the discussion in Part VI below.

Part VI

Addition information on STM/STS measurements and domain switching

6. 1. More discussions on STS spectra at NBC/PBC interface and solitons

The observed soliton is confined at the interface of NBC/PBC chains with a spatial extension of a few nanometers. Fig. S6b presents the intensity line profile taken at in-gap state energy (E = -0.23 eV) in Fig. S6a. It can be fitted using a simple model $I = \frac{I_0}{\xi} \times sech\frac{x-x_0}{\xi \cdot a_0} cos\frac{\pi(x-x_0)}{a_0}$, in which ξ is the coherence length, x_0 is the center position of soliton, and a_0 is the lattice constant of the

unit cell. The extracted coherence length of the soliton decaying into the PBC and NBC chains is found to be ξ_p = 1.01 nm and ξ_n = 1.60 nm, respectively. The overall coherence length of the soliton is about 2.6 nm, or $7a_0$, similar to those observed in polyacetylene (~$7a_0$) [8] and In-4×1/Si(111) (~ $9a_0$) [9].

Aside from the existence of in-gap states, there is another unique feature of soliton - the 180° phase shift in lattice atoms' arrangement across the NBC/PBC boundary, similar to that in polyacetylene [10], which can be reflected in the local density of state (LDOS) measured from STS measurements. Fig. S6(**c, d**) show the STS intensity line profiles taken from the valance and conduction bands (line-cut 3 and line-cut 1, respectively), and both can be fitted by the function $Z = Z_0 + A \times tanh\frac{x-x_0}{\xi \cdot a_0} sin\frac{\pi(x-x_0)}{a_0}$ with a 180° phase shift at the interface, confirming that the Si atom lattice is shifted by half a lattice constant.

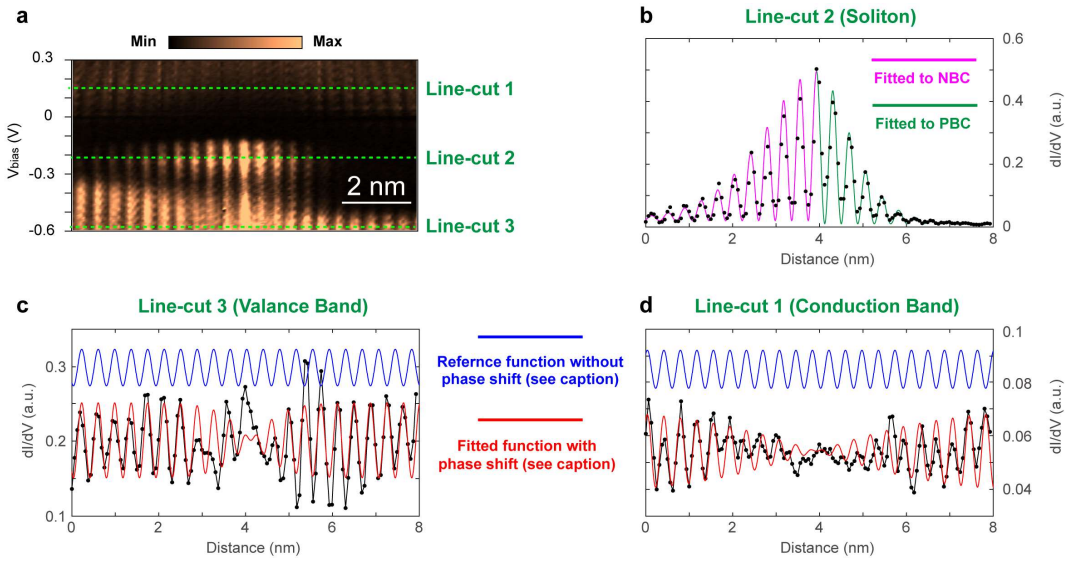


Figure S6 | a, dI/dV spectra across the boundary between PBC and NBC chains (setpoint: V = 0.3 V, I = 20 pA, Osc = 3 meV). **b,** STS intensity line profile taken at the energy of the in-gap state (-0.23 eV, line-cut 2 in **a**). The decay of the STS intensity into NBC and PBC chains can be well fitted (see text) as shown by the purple and green lines, respectively. **c**,**d**, STS intensity line profiles (black) taken at -0.59 eV (line-cut 3 in **a**) and 0.15 eV (line-cut 1 in **a**) with the fitted curves (red) overlapped. The fitted curves have taken into account the 180° phase across the boundary (see text). Blue curve is a sine fuction without the 180° phase shift for comparison - clearly, if the phase is aligned to the NBC chains (left), the phase is mismatched by 180° to the PBC chain.

6.2: Discussion on the switching of NBC/PBC domains

In the main text, we show a series of NBC and PBC domain switching in Fig.4b, c. In Fig. S7, we demonstrate more such results. Before the domain switching (Fig. S7a(i), the surface is mostly coved by PBC domains with a few small NBC patches (marked by red arrows). As the sample bias decreases from Fig. S7a(i) to Fig. S7a(x), the NBC domains grow continuously and eventually dominate the area, leaving only a couple of small PBC patches (see, e.g. Fig. S6a(x)).

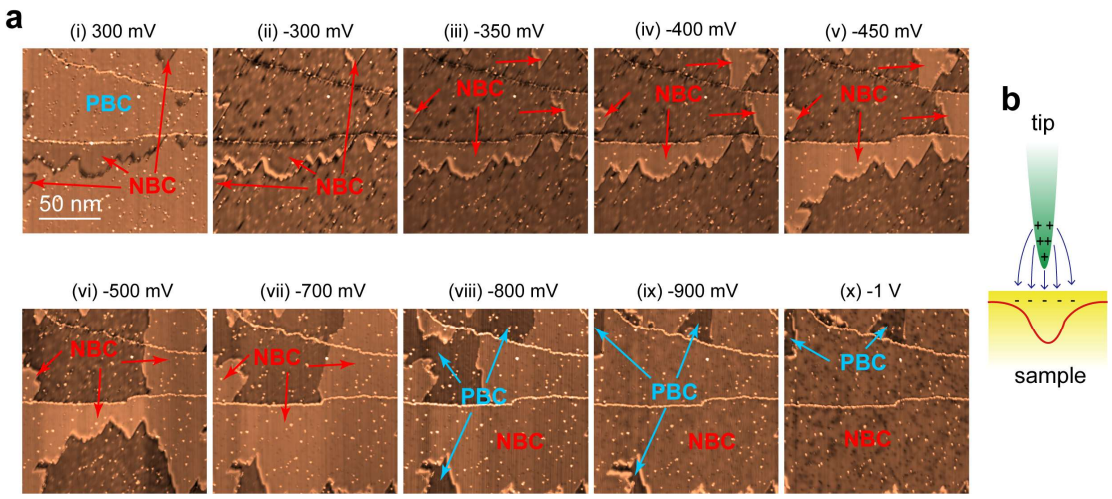


Figure S7 | a, Schematically drawing of the charging effect between the tip and the sample surface. **b**, Sequential topography image of the same area under indicated set points. The scanning current for all images are 40 pA.

Although a detailed investigation on the mechanism of such domain switching is under way (experiments with more controlling parameters are being carried out) and is beyond the scope of this work, we would like to briefly discuss a simple kinetic picture. The switching between PBC and NBC domains depends on their relative energy - which was found to be doping level dependent [11]. The NBC/PBC domains are generally more stable with electron/hole doping, thus with an increasingly negative sample bias from Fig. S7a(i) to Fig. S7a(x), more negative charges will accumulate below the STM tip, as illustrated in Fig. S7b, which favors the NBC domain. Using a planar capacitance model for an order of magnitude estimate, the induced charge density would be on the order of $\varepsilon_0 V/d$. Experimentally, d is the distance between the STM tip and the sample

(~1nm); and the bias V is on the order of 1 V (see Fig. S7). Thus the induced carrier density is ~ 10^{12} cm^{-2}, which is comparable with the intrinsic surface carrier density of Si. Therefore the STM tip could effectively tune the doping level of the local surface atoms and induce the NBC/PBC switching.

6. 4. Determine ('read') NBC/PBC domains using LDOS information.

As plotted in Fig. 4a(iv) and Fig. S8, the dI/dV spectra (reflecting the LDOS) for NBC and PBC domains show a vast difference - up to an order of magnitude (~ 9.4 times, see Fig. S8b(ii)) within the sampe bias range -0.5V ~ -0.3V (see Fig. S8b(ii)). Such a large difference of dI/dV signal can be exploited to discern ("read") the domain type.

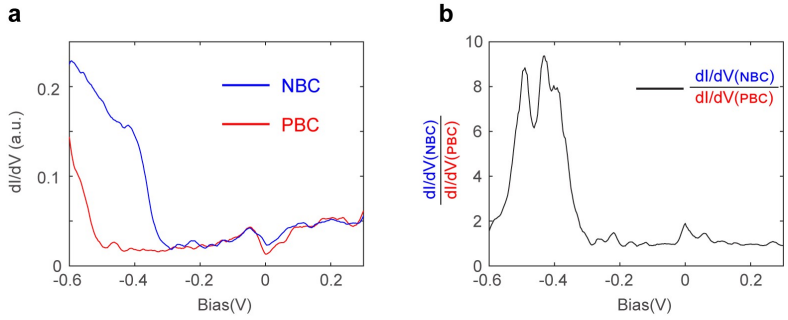


Figure S8 | a, typical dI/dV spectra for NBC and PBC domains. **b**, Bias-dependent dI/dV ratio between NBC and PBC domains.

References:

1. Song, Z.D., L. Elcoro, and B.A. Bernevig, Twisted bulk-boundary correspondence of fragile topology. *Science* **367**, 794-797 (2020).
2. Xu, Y.F. et al., Three-Dimensional Real Space Invariants, Obstructed Atomic Insulators and A New Principle for Active Catalytic Sites. *submitted* (2021).
3. Kresse, G. and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Physical Review B* **54**, 11169-11186 (1996).
4. Blöchl, P.E., Projector augmented-wave method. *Physical Review B* **50**, 17953-17979 (1994).

5. Bussetti, G., et al., Coexistence of Negatively and Positively Buckled Isomers on n^+-Doped Si(111) - 2 x 1. *Phys. Rev. Lett.* **106**, 067601 (2011).

6. Loser, K., et al., Spectroscopy of positively and negatively buckled domains on Si(111)-2x1. *Physical Review B* **86**, 085303 (2012).

7. Violante, C., et al., Geometric, electronic, and optical properties of the Si(111)2x1 surface: Positive and negative buckling. *Physical Review B* **86**, 245313 (2012).

8. Su, W.P., J.R. Schrieffer, and A.J. Heeger, Solitons in Polyacetylene. *Physical Review Letters* **42**, 1698-1701 (1979).

9. Kim, T.H. and H.W. Yeom, Topological Solitons versus Nonsolitonic Phase Defects in a Quasi-One-Dimensional Charge-Density Wave. *Physical Review Letters* **109**, 246802 (2012).

10. Heeger, A.J., et al., Solitons in conducting polymers. *Reviews of Modern Physics* **60**, 781-850 (1988).

11. Feenstra, R.M., et al., Charge transfer between isomer domains on n^+-doped Si(111)-2 x 1: energetic stabilization. *Journal of Physics-Condensed Matter* ***24***, 354009 (2012).