1. (a) All entropy formulae below have an arbitrary multiplicative constant $K$ setting the units, or alternatively an arbitrary base of the logarithm.
joint information entropy:

$$
H(X, Y)=-\sum_{i, j} p_{i j} \log p_{i j}
$$

conditional information entropies:

$$
\begin{align*}
& H(X \mid Y)=\sum_{j} p_{j}^{(Y)} H\left(X \mid Y=y_{j}\right)=-\sum_{i, j} p_{i j} \log \frac{p_{i j}}{p_{j}^{(Y)}} \\
& H(Y \mid X)=\sum_{i} p_{i}^{(X)} H\left(Y \mid X=x_{i}\right)=-\sum_{i, j} p_{i j} \log \frac{p_{i j}}{p_{i}^{(X)}}
\end{align*}
$$

mutual information:

$$
I(X ; Y)=\sum_{i, j} p_{i j} \log \frac{p_{i j}}{p_{i}^{(X)} p_{j}^{(Y)}}
$$

## [Bookwork]

(b) Three independent relations:

$$
\begin{gathered}
H(X \mid Y)=H(X, Y)-H(Y) \\
H(Y \mid X)=H(X, Y)-H(X) \\
I(X ; Y)=H(X)+H(Y)-H(X, Y)
\end{gathered}
$$

any set of independent relations worth $\{\mathbf{2}\}$ per equation.
[Bookwork]
(c) Using notation $\{e, o\}$ for even, odd (for $V$ ) and $\{p, n\}$ for prime, non-prime (for $W$ ), the joint probabilities: $p_{e p}=1 / 6, p_{e n}=1 / 3, p_{o p}=1 / 3, p_{o n}=1 / 6$. Using bits as units (using $\log _{2}$ ), the joint entropy:

$$
\begin{align*}
H(V, W) & =-\left(\frac{1}{6} \log _{2} \frac{1}{6}+\frac{1}{3} \log _{2} \frac{1}{3}+\frac{1}{3} \log _{2} \frac{1}{3}+\frac{1}{6} \log _{2} \frac{1}{6}\right) \\
& =\frac{1}{3} \log _{2} 6+\frac{2}{3} \log _{2} 3=\log _{2} 3+\frac{1}{3} \\
& =\frac{\ln 3}{\ln 2}+\frac{1}{3} \approx 1.92 \text { bit }
\end{align*}
$$

One way to calculate the mutual information is to realise that $p_{e}=p_{o}=1 / 2$, so $H(V)=-2 \frac{1}{2} \log _{2} \frac{1}{2}=1$ bit, similarly $H(W)=1$ bit.
Then

$$
\begin{align*}
I(V ; W) & =H(V)+H(W)-H(V, W) \\
& =1+1-\left(\log _{2} 3+\frac{1}{3}\right)=\frac{5}{3}-\log _{2} 3 \\
& \approx 0.08 \mathrm{bit}
\end{align*}
$$

Alternatively, one can apply the formula in 1(a).
[Unseen]
(d) (i) False. They don't even have the same units (eg. bit vs bit ${ }^{2}$ )!

A numerical counterexample (using bits as units): if $X$ and $Y$ are independent coin tosses, $H(X)=H(Y)=1, H(X, Y)=2 \neq 1 \cdot 1$.
(ii) False. It can be easily shown that $H(X, X)=H(X)$ (eg. by applying the definition). Then $H(X \mid X)=H(X, X)-H(X)=0$, so any nontrivial $X$ is counterexample.
(iii) True. $I(X ; X)=H(X)+H(X)-H(X, X)=H(X)$.
[Unseen]
2. (a) Interface width:

$$
w(L, t)=\sqrt{\left\langle(h(x, t)-\bar{h}(t))^{2}\right\rangle_{x}} \quad \text { where } \quad \bar{h}(t)=\langle h(x, t)\rangle_{x}
$$

## [Bookwork]

(b) (i) early times, $t \ll t_{\times}: \quad w(L, t) \sim t^{\beta}$ late times, $\quad t \gg t_{\times}: \quad w(L, t) \sim w_{\text {sat }}(L) \sim L^{\alpha}$ $\beta$ : growth exponent crossover time: $\quad t_{\times} \sim L^{z} \quad z$ : dynamic exponent $\{\mathbf{2}\}$

## [Bookwork]

$\alpha$ : roughness exponent
(ii) Family-Vicsek scaling relation:

$$
w(L, t) \sim L^{\alpha} f\left(\frac{t}{L^{z}}\right) \quad f(u) \sim \begin{cases}u^{\beta}, & \text { if } u \ll 1 \\ \text { const, } & \text { if } u \gg 1\end{cases}
$$

For $t \ll L^{z}: W(L, t) \sim L^{\alpha} \cdot\left(t / L^{z}\right)^{\beta} \sim t^{\beta}($ using $z=\alpha / \beta)$
For $t \gg L^{z}: W(L, t) \sim L^{\alpha}$. const $\sim L^{\alpha}$.

## [Bookwork]

(iii) To recover $t^{\beta}$ for small $t$, we need $g(u)=$ const for $u \gg 1$ (assuming $B>0$ ) to make it independent of $L$, which sets $A=\beta$.
To obtain $L^{\alpha}$ for large $t$, we need $g(u)=u^{\alpha}$ for $u \ll 1$.
Then for large $t, L^{\alpha} \sim w \sim t^{\beta} L^{\alpha} t^{-B \alpha}$. This gives $B=\beta / \alpha=1 / z$.
In summary:

$$
w(L, t) \sim t^{\beta} g\left(\frac{L}{t^{1 / z}}\right) \quad g(u) \sim \begin{cases}u^{\alpha}, & \text { if } u \ll 1 \\ \text { const, } & \text { if } u \gg 1\end{cases}
$$

[Unseen]
(c) (i) Random deposition model: the interface grows on a discretised substrate by accreting squares (hypercubes in general dimension). The cubes arrive above random substrate positions, and simply increase the height of the colum at that position.

## [Bookwork]

(ii) Suppose $N$ cubes are grown on a substrate made of $S$ units. (In one dimension $S=L$, in two dimensions $S=L^{2}$ etc.) The growth above a given substrate location can be considered as the sum of $N$ i.i.d Bernoulli processes, each grows unit height $(\Delta h=1)$ with probability $p=1 / S$ and does not grow $(\Delta h=0)$ with probability $1-p$. The expectation $\langle\Delta h\rangle=p=1 / S$, and the variance $\operatorname{Var}(\Delta h)=\left\langle(\Delta h)^{2}\right\rangle-\langle\Delta h\rangle^{2}=p(1-p)=(S-1) / S^{2}$, since $\left\langle(\Delta h)^{2}\right\rangle=p$

The Central Limit Theorem says that the sum of $N$ such processes has mean $\bar{h}=$ $N\langle\Delta h\rangle=N / S$, and variance

$$
\left\langle(h-\bar{h})^{2}\right\rangle=N \operatorname{Var}(\Delta h)=N \frac{S-1}{S^{2}}=w^{2}
$$

which equals to the square of the interface width. Now fix the substrate size $S$. If time is measured as $N$ (maybe with some proportionality constant, eg. $t=N / S$; this does not change the scaling exponent), then

$$
w \sim \sqrt{N} \sim \sqrt{t}
$$

using $w \sim t^{\beta}$, this gives $\beta=1 / 2$.
[Unseen]
(iii)

$$
\frac{\partial h}{\partial t}=F+\eta(x, t)
$$

Where $\eta$ has zero mean and is delta correlated:

$$
\begin{gathered}
\langle\eta(x, t)\rangle=0 \\
\left\langle\eta(x, t) \eta\left(x^{\prime}, t^{\prime}\right)\right\rangle=2 D \delta\left(x-x^{\prime}\right) \delta\left(t-t^{\prime}\right)
\end{gathered}
$$

[Bookwork]
3. (a) The laws of thermodynamics (credit $\{\mathbf{1}\}$ for each):
(0) There exist a relation between thermodynamic systems. This relation is called thermodynamic equilibrium, and it is transitive (equivalence relation):
if $A \sim B$ and $B \sim C$, then $A \sim C$.
Here $A, B$ and $C$ label different systems. For example in thermal equilibrium this means a transitive relation between the temperatures of the three systems.
(1) Energy conservation: the total energy of an isolated system is fixed.

Thus if during some process a system absorbs heat $\Delta Q$, as well as work $\Delta W=$ $-p \Delta V+\ldots$ is made on it, then its energy changes by $\Delta E=\Delta Q+\Delta W$.
(2) In an isolated system the entropy does not decrease.

Thus if during some process a system absorbs heat $\Delta Q$, then its entropy changes by $\Delta S=\Delta Q / T+\Delta S_{\text {internal }} \geq \Delta Q / T$.
(3) The entropy at absolute zero temperature is zero (or can be set zero).

## [Bookwork]

(b) (i) Free energies are the Legendre transforms of the energy.
(ii) The free energy of a system does not increase, so at stable equilibrium it is minimal.

An example: (canonical ensemble:) a system is kept at fixed temperature $T$ while undergoing some change.
Its change in energy: $\Delta E=\Delta Q$.
Its change in entropy: $\Delta S=\Delta Q / T+\Delta S_{\text {internal }} \geq \Delta Q / T$
The relevant free energy is the Helmholtz free energy: $A(T)=E-T S$.
The change in Helmholtz free energy: $\Delta A=\underbrace{\Delta E}_{\Delta Q}-\underbrace{T \Delta S}_{\geq \Delta Q} \leq 0$.
(iii) The probability of a macroscopic state (sum of Boltzmann factors) can be expressed by a single Boltzmann factor, in which the energy is replaced by the appropriate free energy.
Example: in a grand canonical ensemble the probabilities multiplied by $\Xi$ :

$$
\Xi=\sum_{i} e^{-\beta\left(E_{i}-\mu N_{i}\right)}=\sum_{N=0}^{\infty} e^{\beta \mu N} \underbrace{\sum_{j} e^{-\beta E_{j ; N}}}_{\text {state with same } N}=\sum_{N=0}^{\infty} e^{-\beta(A(T ; N)-\mu N)}
$$

[Bookwork]
(c) (i) The partition function:

$$
\begin{align*}
Z= & \frac{1}{h^{2}} \int_{-\infty}^{\infty} d x \int_{-\infty}^{\infty} d y \int_{-\infty}^{\infty} d p_{x} \int_{-\infty}^{\infty} d p_{y} \\
& \times \frac{\exp \left[-\beta\left(\frac{p_{x}^{2}}{2 m}+\frac{p_{y}^{2}}{2 m}+\frac{m g}{2 \ell}\left(x^{2}+y^{2}\right)\right)\right]}{h^{2}} \underbrace{\int_{-\infty}^{\infty} d x e^{-\beta \frac{m g}{2 \ell} x^{2}} \int_{-\infty}^{\infty} d y e^{-\beta \frac{m g}{2 \ell} y^{2}}}_{\sqrt{\frac{2 \pi \ell}{\beta m g}}} \\
& \times \underbrace{\int_{-\infty}^{\infty} d p_{x} e^{-\beta \frac{1}{2 m} p_{x}^{2}}}_{\sqrt{\frac{2 \pi m}{\beta}}} \int_{-\infty}^{\infty} d p_{y} e^{-\beta \frac{1}{2 m} p_{y}^{2}} \\
= & \frac{\ell}{\hbar^{2} g \beta^{2}}
\end{align*}
$$

(ii) Average energy:

$$
\langle E\rangle=-\frac{\partial \ln Z}{\partial \beta}=-\frac{\partial}{\partial \beta} \ln \frac{1}{\beta^{2}}=\frac{2}{\beta}=2 k_{B} T
$$

(iii) Each quadratic half-degree of freedom (eg. $\frac{p_{x}^{2}}{2 m}$ or $\frac{m g x^{2}}{2 \ell}$ ) contributes $\frac{1}{2} k_{B} T$ to the average energy.
In this case we have 4 such quadratic contributions, leading to $\langle E\rangle=2 k_{B} T . \quad\{\mathbf{1}\}$
[Unseen] in this form, though harmonic oscillator was covered in lecture.

