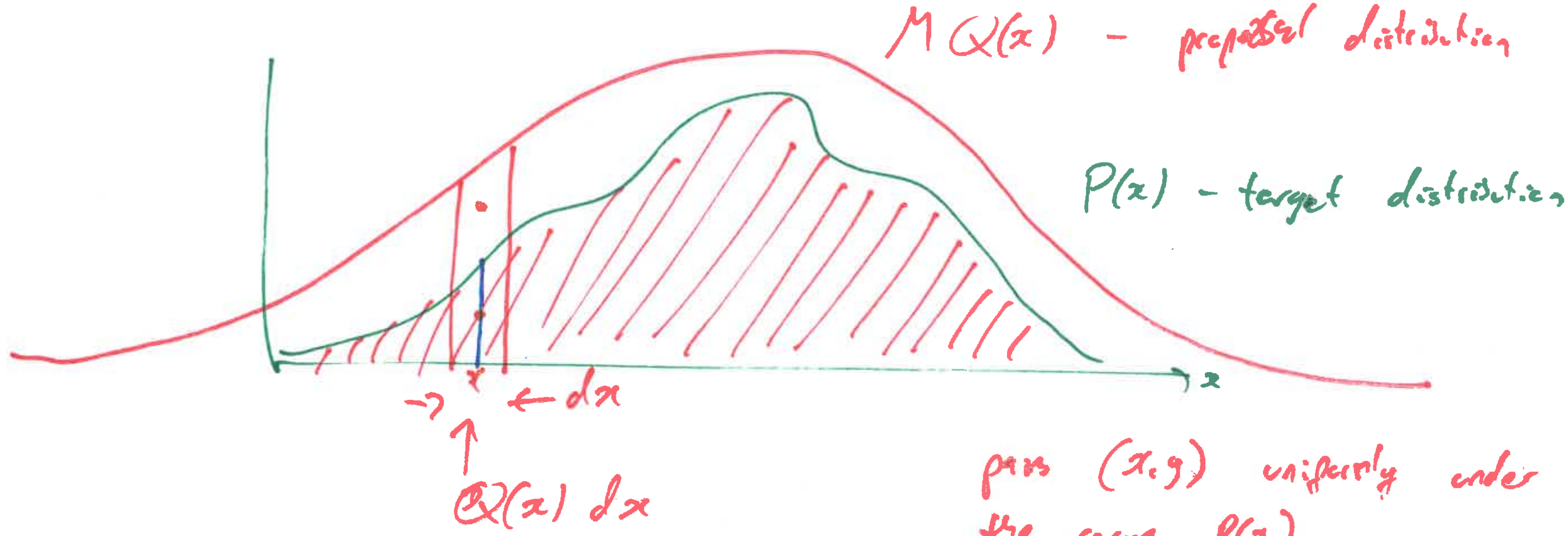
Slice Sampling

One final example of an MCMC algorithm. (Neal 2003).

Inspired by rejection sampling. (1-dimensional case $\mathcal{X} = \mathbb{R}$.)

Recall Rejection sampling

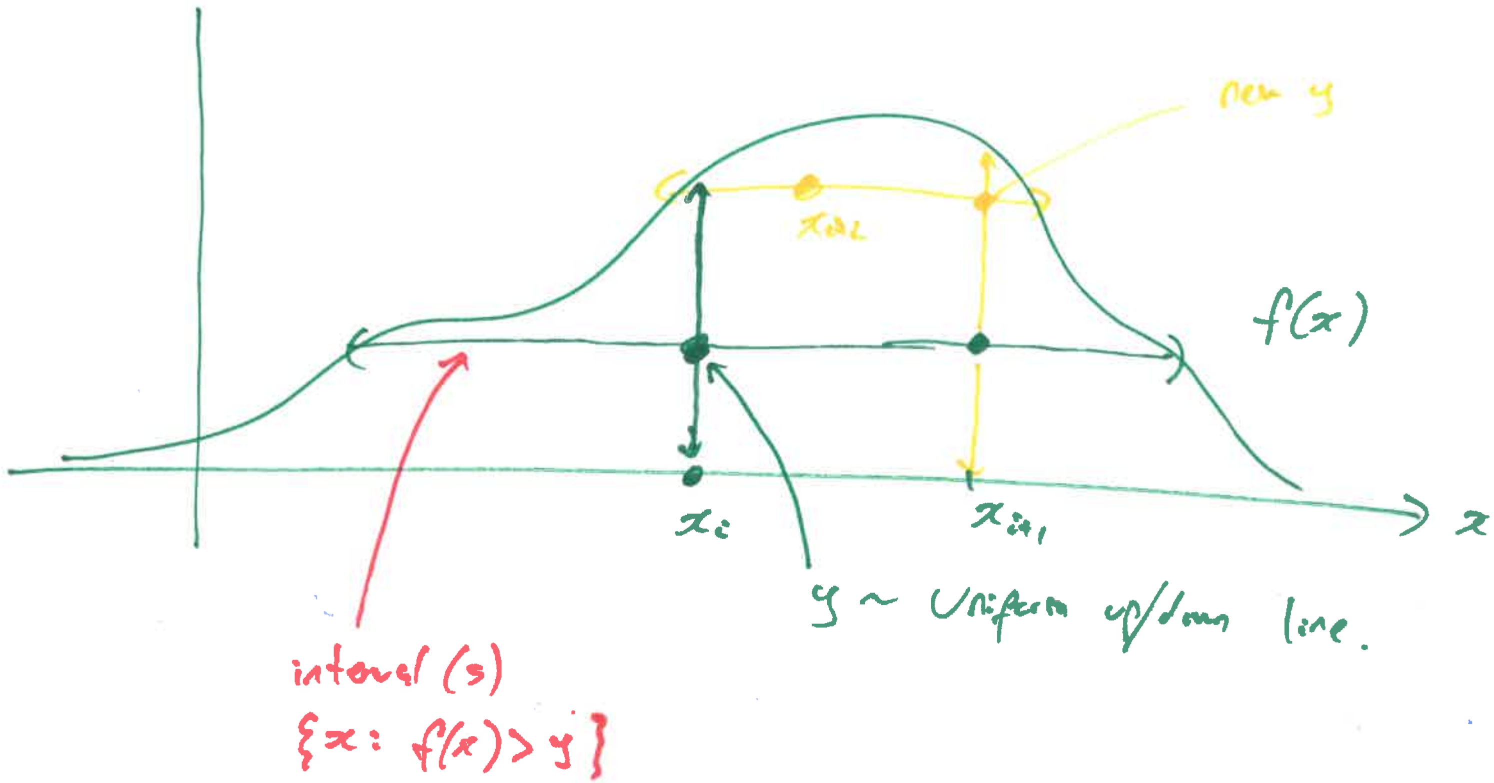


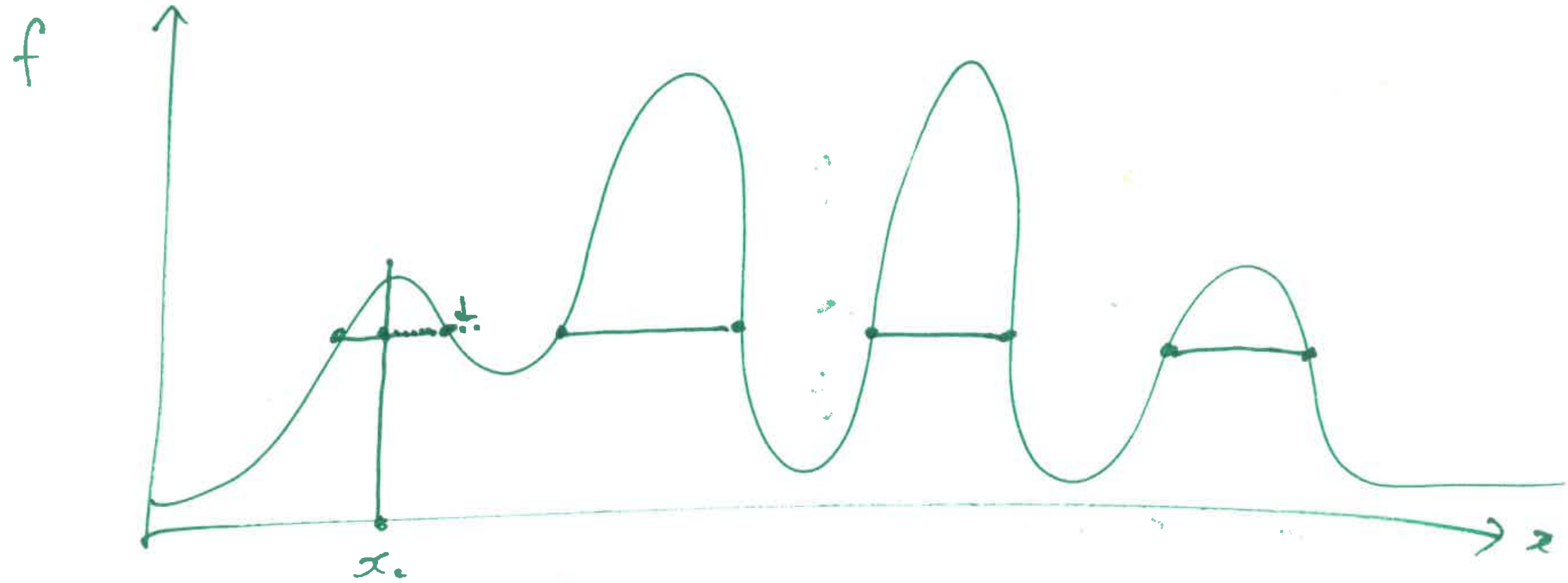
pairs (x, y) uniformly under the curve $P(x)$.
discard y to leave $\boxed{x \sim P(x)}$

Slice sampling $f(x) \propto P(x)$

<u>Alogrithm</u>:

1. $x_0 \sim \alpha$ Initialise MC
2. $i \leftarrow 0$
3. while $i \geq 0$ do iterating $i = 0, 1, 2, \ldots$
4. $y \sim U(0, f(x_i))$ ~~auxiliary~~ auxiliary variable.
5. $x_{i+1} \sim U(\{x : f(x) > y\})$
6. end while.


new y
x_{i+2}
$f(x)$
x
x_i
x_{i+1}
$y \sim$ Uniform up/down line.
interval (s)
$\{x : f(x) > y\}$


f
t_i
x_0
x

Lemma: Slice sampling will produce MC $x_0, x_1, x_2, \ldots$ will have as its stationary dist. $P(x)$.

The alg. uses $P(x) = \frac{f(x)}{Z}$.

Proof. let's suppose $x_i \sim P$

$y \mid x_i$ ~~y~~ $\sim U(0, f(x_i))$

$x_{i+1} \mid y \sim U(\{x_{i+1} : f(x_{i+1}) > y\})$

$$\int dy \int dx_i \, \underbrace{P(x_{i+1} \mid y)} \, \underbrace{P(y \mid x_i)} \, P(x_i) \stackrel{?}{=} P(x_{i+1})$$

$$P(x_{i+1} \mid y) = \frac{\mathbb{1}_{\{x_{i+1} : f(x_{i+1}) > y\}}(x_{i+1})}{\int_{\{z : f(z) > y\}} dz \cdot 1}$$

$$P(y \mid x_i) = \frac{\mathbb{1}_{(0, f(x_i))}(y)}{f(x_i)}$$

$$\int dy \int dx_i \; \frac{\mathbb{1}_{\{x_{i+1}:\, f(x_{i+1})>y\}}(x_{i+1})}{\int_{\{z:\, f(z)>y\}} dz \cdot 1} \; \frac{\mathbb{1}_{(0,f(x_i))}(y)}{f(x_i)} P(x_i)$$

$$\frac{1}{Z} \int dy \; \frac{\mathbb{1}_{\{x_{i+1}:\, f(x_{i+1})>y\}}(x_{i+1})}{\int_{\{z:\, f(z)>y\}} dz} \int dx_i \; \mathbb{1}_{(0,f(x_i))}(y) = 1$$

$$\frac{1}{Z} \int dy \; \mathbb{1}_{\{x_{i+1}:\, f(x_{i+1})>y\}}(x_{i+1})$$

$$\frac{1}{Z} \; f(x_{i+1})$$

$$P(x_{i+1})$$

$\square$

Slice it's usually only implemented in 1D as part of a Gibbs sampling algorithm.

$$P(x^0 \mid x^1, x^2, x^3 \ldots x^{d-1})$$

sample this using slice sampling.

In this way "Slice within Gibbs" sample can tackle high-dim. problems.

Neal 2003

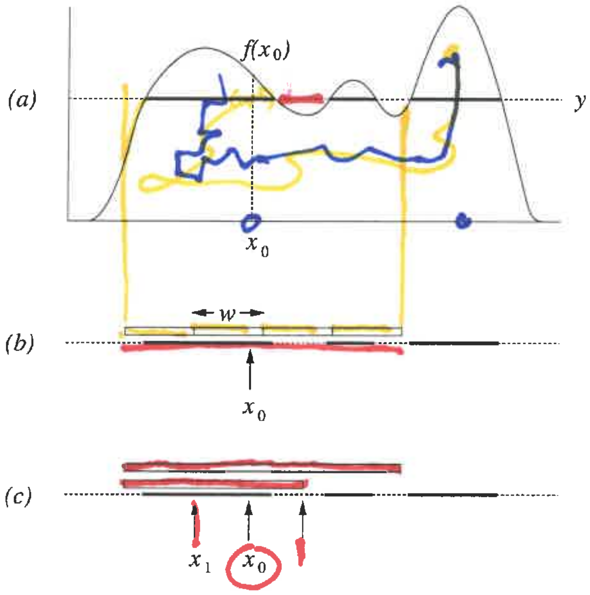


FIG. 1. *A single-variable slice sampling update using the stepping-out and shrinkage procedures. A new point, x_1, is selected to follow the current point, x_0, in three steps. (a) A vertical level, y, is drawn uniformly from $(0, f(x_0))$, and used to define a horizontal 'slice', indicated in bold. (b) An interval of width w is randomly positioned around x_0, and then expanded in steps of size w until both ends are outside the slice. (c) A new point, x_1, is found by picking uniformly from the interval until a point inside the slice is found. Points picked that are outside the slice are used to shrink the interval.*

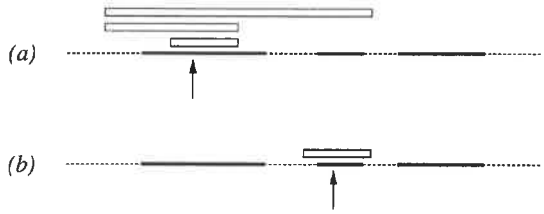


FIG. 2. *The doubling procedure. In (a), the initial interval is doubled twice, until both ends are outside the slice. In (b), where the start state is different, and the initial interval's ends are already outside the slice, no doubling is done.*

Steps *(b)* and *(c)* can potentially be implemented in several ways, which must of course be such that the resulting Markov chain leaves the distribution defined by $f(x)$ invariant. Figure 1 illustrates one generally-applicable method, in which the interval is found by 'stepping out', and the new point is drawn with a 'shrinkage' procedure. Figure 2 illustrates an alternative 'doubling' procedure for finding the interval. These and some other variations are described in detail next, followed by a proof that the resulting transitions leave the correct distribution invariant.