# EWIIS3 2021

## ① Consumption tariff

MDP $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{R}, \gamma)$

State space $\mathcal{S} = A \times B = \{ (\alpha, \beta) \mid \alpha \in A, \beta \in B \}$

MUBP = mean usage-based price

MUBP(c) = average over all rates specified in a given consumption tariff

$MUBP_{min} = \min_{c} MUBP(c)$

$c \in \mathcal{C} = \{ c_0, c_1, \dots, c_{n-1} \}$ (active consumption tariffs)

c of EWIIS3: $c^{EWIIS3}$

$\alpha$ = percentual deviation of EWIIS3's MUBP from active consumption tariff

with lowest MUBP: $\alpha = \dfrac{MUBP_{min} - MUBP(c^{EWIIS3})}{|MUBP_{min}|}$

In EWIIS3 2020: discrete "MUBP status" (MUBPS)

$\beta$ = periodic payment factor (PPF)

In EWIIS3 2020: PPF $\in \{4, 6, 8, 10\}$

Here: e.g., $\underbrace{\beta \in [0, 10]}_{B}$, same meaning as PPF

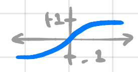
state $s = (\alpha, \beta)$

Action space $\mathcal{A} = \{ (\alpha', \beta') \mid \alpha' \in A, \beta' \in B \}$

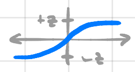
broker decides on how great it wants percentual dev. to $MUBP_{min}$ to be (?)

e.g. broker only decides on difference $\Delta\alpha$ or factor $\lambda_\alpha$:

$\alpha' = \underbrace{z \cdot \tanh(f(s))}_{\Delta\alpha} + \alpha$ $\qquad \tanh(\cdot) \in [-1, +1]$



$z \cdot \tanh(\cdot) \in [-z, +z]$



$\alpha' = \underbrace{z \cdot \tanh(f(s))}_{\lambda_\alpha} \cdot \alpha$

output of pre-activation neural network layer

Same transformation into time-of-use price scheme as before

except using $\beta$ instead of PPF.

e.g. looking at share factor $s_t$

Whether to start a new iteration: not hard coded but decided

by broker in each timeslot $t$: $p_t \in [0, 1]$. New iteration if $p_t > 0.5$.

No more consideration of residential share factors $s_t^{Brookside}$, $s_t^{Centerville}$

Reward: $r_t = \begin{cases} \text{penalty} \ \text{(e.g. } -100{,}000\text{)} & \text{if penalizeRevoke} \\ \boxed{\sum_{i=t-\tau}^{t} (u_i + p_i) - fees_i} & \text{else} \end{cases}$

or: $r_t = \begin{cases} \text{penalty} & \text{if penalizeRevoke} \\ 0 & \text{else} \end{cases}$

for all $t < T$, and only final reward = final cash position $(r_T)$

$u_i$: usage-based charges

$p_i$: periodic charges

## ② Wholesale / balancing markets

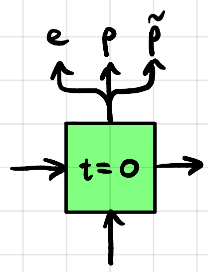
Action space: $\{ (t, e, p, \tilde{p}) \mid t \in \{1, \dots, 24\},$ — timeslot

$e \in ?,$ — electricity amount

$p \in ?,$ — limit price

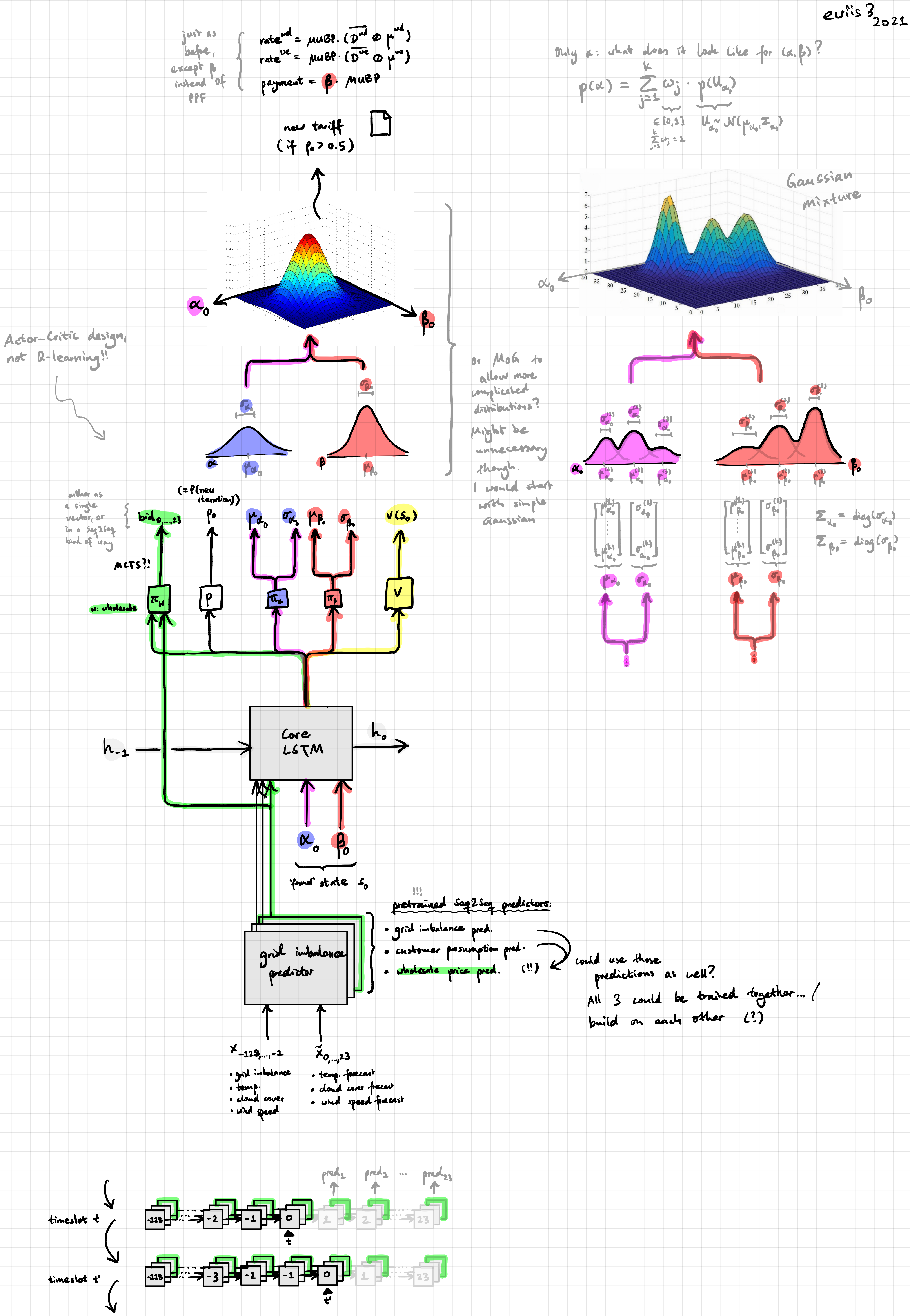
$\tilde{p} \in [0, 1] \}$ — whether to do limit order (0), market order (1)

Can be done recurrently:




How to integrate into rest of broker?

Use MCTS? What would that look like?



eu'is 3 2021
just as before, except β instead of PPF
rate^wd = MUBP · (D^wd ⊘ μ^wd)
rate^ue = MUBP · (D^ue ⊘ μ^ue)
payment = β · MUBP
new tariff (if p₀ > 0.5)
Only α: what does it look like for (α, β)?
p(α) = Σ_{j=1}^{k} ω_j · p(U_{α₀})
∈ [0,1]
Σ_{j=1}^{k} ω_j = 1
U_{α₀} ~ N(μ_{α₀}, Σ_{α₀})
Gaussian mixture
α₀
β₀
Actor-Critic design, not Q-learning!!
or MoG to allow more complicated distributions?
Might be unnecessary though.
I would start with simple Gaussian
Σ_{α₀} = diag(σ_{α₀})
Σ_{β₀} = diag(σ_{β₀})
either as a single vector, or in a Seq2Seq kind of way
(= P(new iteration))
bid_{0,...,23}
p₀
μ_{α₀}
σ_{α₀}
μ_{β₀}
σ_{β₀}
v(s₀)
MCTS?!
w: wholesale
π_w
P
π_α
π_β
V
h₋₁
Core LSTM
h₀
"formal" state s₀
!!!
pretrained Seq2Seq predictors:
• grid imbalance pred.
• customer prosumption pred.
• wholesale price pred. (!!)
could use those predictions as well?
All 3 could be trained together... / build on each other (?)
grid imbalance predictor
x_{-123,...,-1}
• grid imbalance
• temp.
• cloud cover
• wind speed
x̃_{0,...,23}
• temp. forecast
• cloud cover forecast
• wind speed forecast
timeslot t
timeslot t'
pred₁
pred₂
pred₂₃