Combining physical modeling and additive synthesis as a mapping strategy for realtime control

Philippe Guillemain
Laboratoire de Mécanique et d'Acoustique
CNRS-LMA

Vincent Verfaille
Sound Processing & Control Lab. (SPCL)
Input Devices & Music Interaction Lab. (IDMIL)

♪ ICMC07 – Copenhagen, Denmark – August 29, 2007

Introduction

Description of the synthesis models used

Physically-controlled synthesis

Discussion around an example

Conclusions

Introduction

Description of the synthesis models used

Physically-controlled synthesis

Discussion around an example

Conclusions

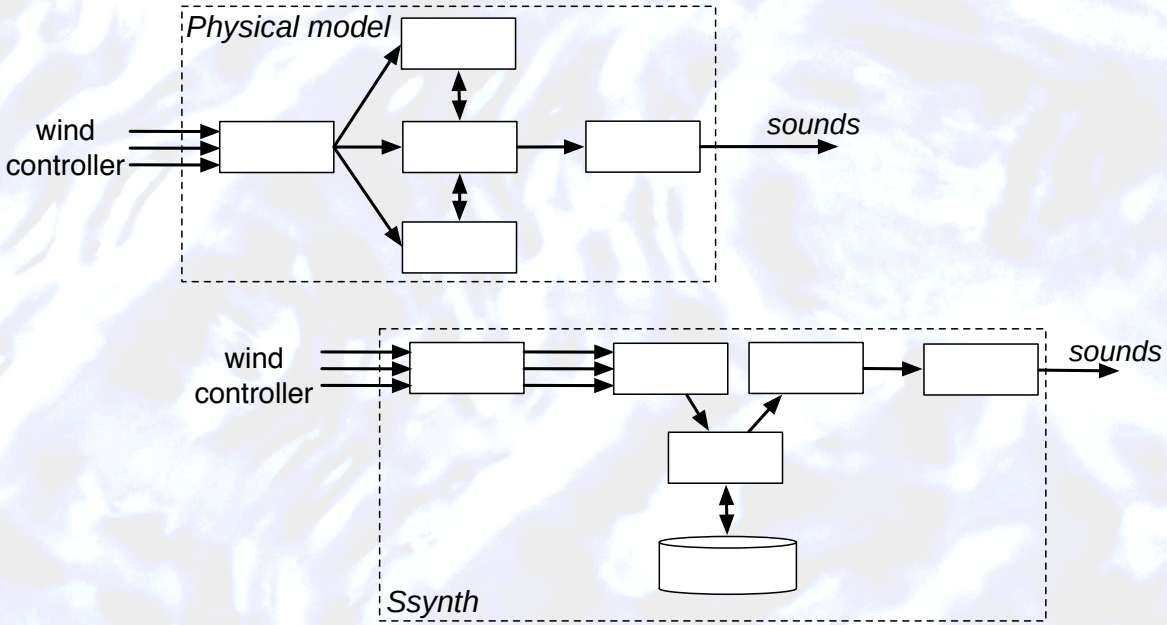
Context

- *Consonnes* project: control of natural and synthetic sounds
- Context: control of self-sustained instruments
  $\Longrightarrow$ woodwinds, in particular clarinet
  challenging problem: nonlinear functioning
- Goal: to build a synthesizer that
  - obeys to physical controls
  - has great sound quality
  - provides mapping/interpolation options
- Case study: clarinet model(s) + WX5 wind controller

Introduction

Technique: interface

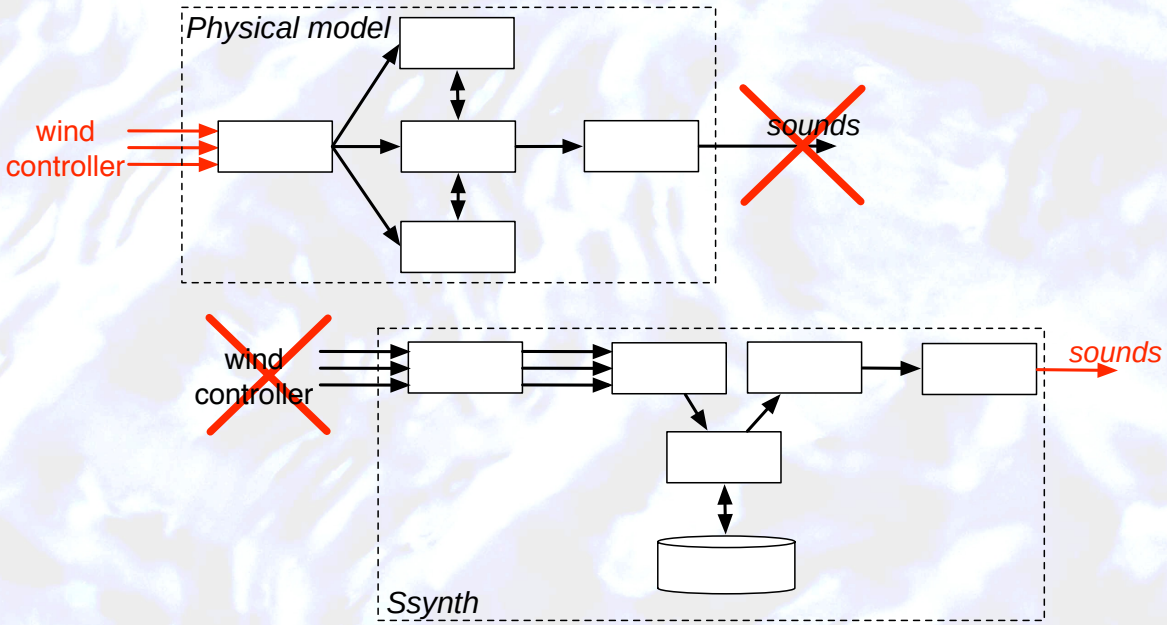
- a physical model with
- an additive re-synthesizer



Introduction

We want to keep:

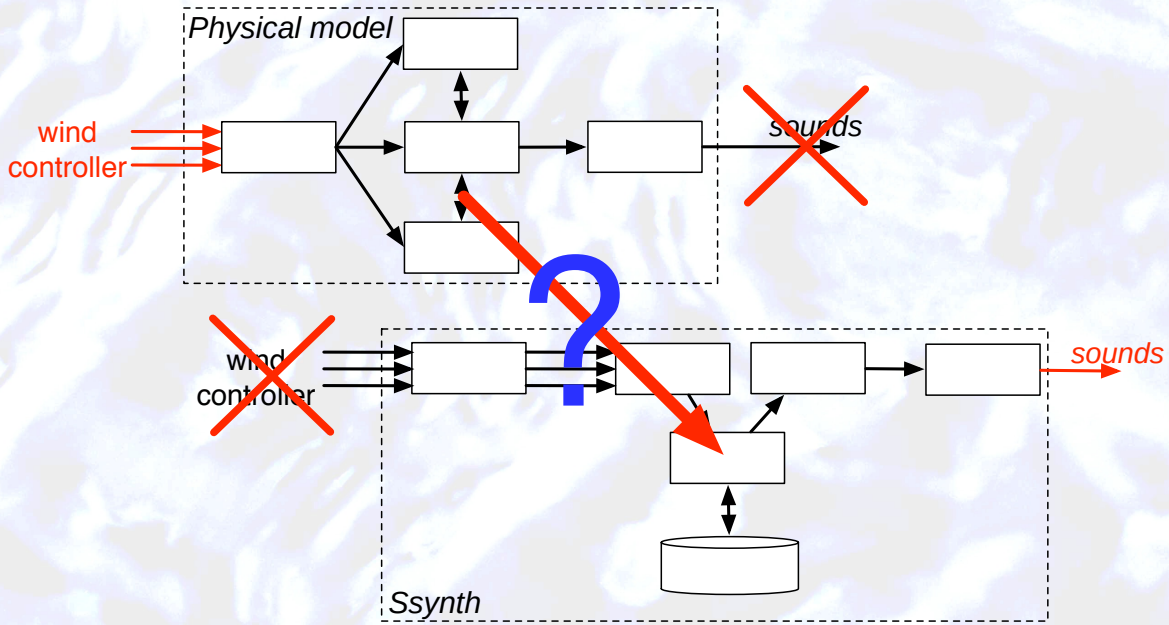
- PM: realistic control ability of physical controls
- AS: sound quality, mapping (interpolation, design)



Introduction

Question: how & where do we interface them?

1. best for realtime & quality: spectrum analytical description in the (too simple) physical model *[Kergomard* et al., *2000]*

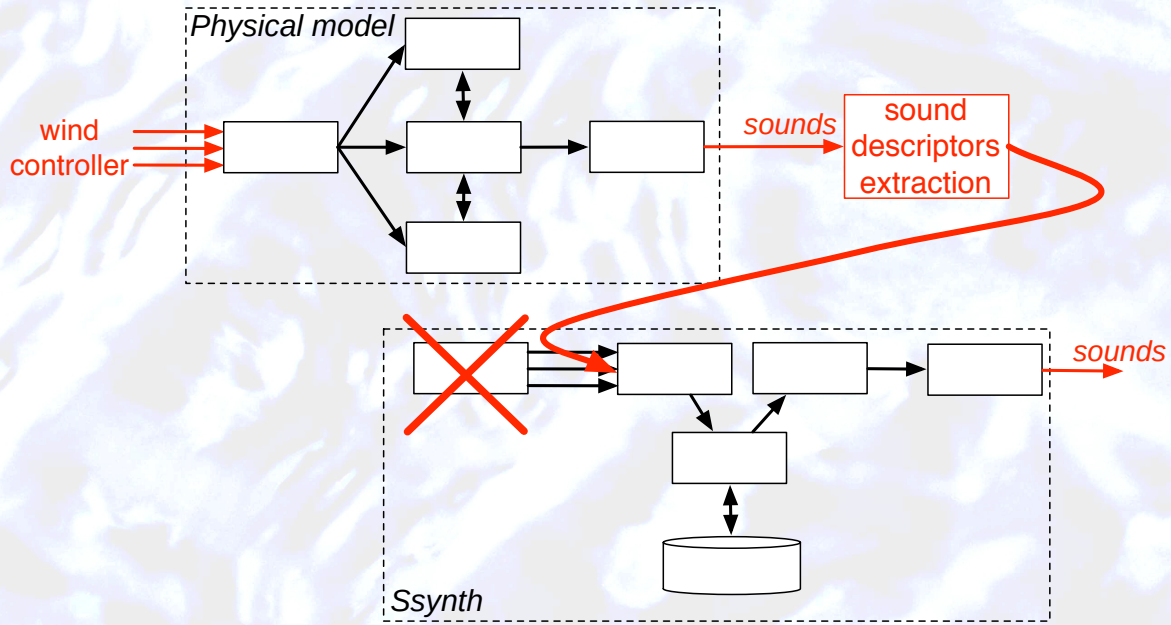


Kergomard, Ollivier, Gilbert, “Calculation of the spectrum of self-sustained oscillators using a variable truncation method: Application to cylindrical reed instruments,” *Acustica* (86), pp. 685–703, 2000.

Introduction

Question: how & where do we interface them?

2. proof of concept: quick and dirty
   $\Longrightarrow$ use sound descriptors, ie. $\sim$ audio-driven



Introduction

Description of the synthesis models used

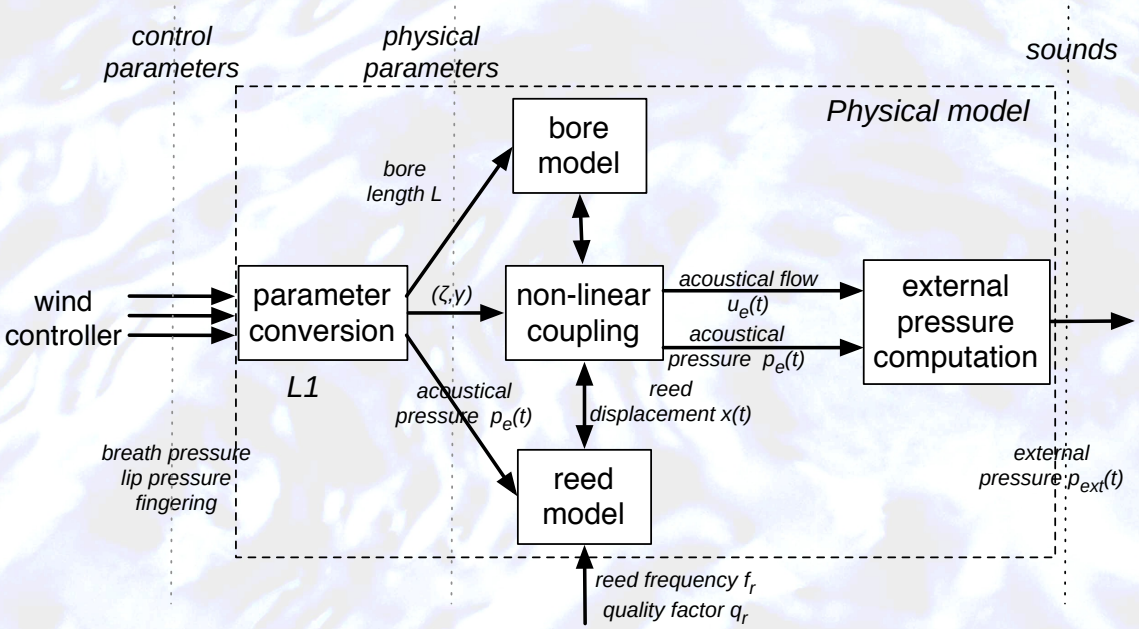
Physically-controlled synthesis

Discussion around an example

Conclusions

Simplified physical model of the clarinet

Model = non-linear coupling of the bore with the reed



Rmq: adimensional variables as in:

Guillemain, Kergomard, Voinier, “Real-time synthesis of clarinet-like instruments using digital impedance models,” JASA (118)1, pp. 483–494, 2005.

Simplified physical model of the clarinet

[more in the paper]

1. bore model:
   - perfect cylinder of length L
   - mouthpiece input : $Z_e(\omega) = \frac{P_e(\omega)}{U_e(\omega)} = i \tan(k(\omega)L)$
   - $k(\omega)$: visco-thermal losses approximation

   $\Longrightarrow$ $\forall$ flow U_e, $\exists$ mainly odd harmonics in P_e (acous. pressure)
2. reed model: pressure driven mass-spring oscillator

   $\frac{1}{\omega_r^2}\frac{d^2x(t)}{dt^2} + \frac{q_r}{\omega_r}\frac{dx(t)}{dt} + x(t) = p_e(t)$

   $x(t)$: reed displacement under $p_e(t)$
3. non-linearity: steady Bernoulli equation

   reed channel opening: $S(t) = \Theta(1 - \gamma + x(t))\zeta[1 - \gamma + x(t)]$

   Θ: Heaviside fonction

Simplified physical model of the clarinet

3. non-linearity:
   - mouthpiece: $\zeta \in [0.2, 0.6]$
     (lip position + section ratio of mouthpiece opening / resonator)
   - reed position: $\gamma \in \left[\frac{1}{3}, 1\right]$
     (oscillation / beating / extinction threshold)
   - acoustical flow:
     $u_e(t) = S(t)\text{sign}(\gamma - p_e(t))\sqrt{|\gamma - p_e(t)|}$
     $\Longrightarrow$ explains even harmonics
4. external pressure: $p_{ext}(t) = \frac{d}{dt}(p_e(t) + u_e(t))$

$\Longrightarrow$ odd/even balance: signature of the non-linearity strength

Rmk: behaviours observed on real and synthetic sounds:

Guillemain, Helland, Kronland-Martinet, Ystad, "The clarinet timbre as an attribute of expressiveness," *Lecture Notes on Computer Science* (3310), pp. 246–259, 2004.

Helland, "Synthesis models as a tool for timbre studies," Master's thesis, 2004.

Additive re-synthesis with *Ssynth*

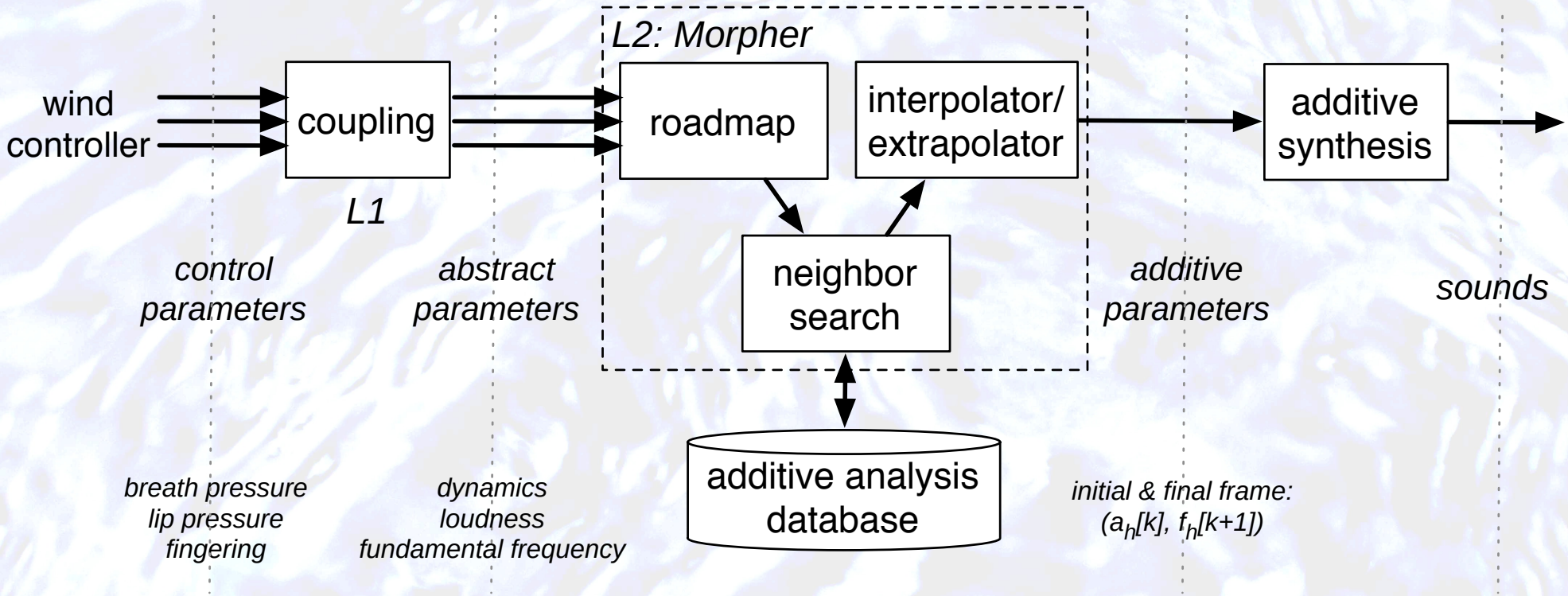
[more in the paper]

- *Ssynth* *[Verfaille* et al.*, 2006]*:
  - real time additive (re)synthesizer
  - advanced and flexible control functionalities
  - developed for studying interpolation of digital musical instruments playing
  - successor of *Escher* *[Wanderley* et al.*, 1998]*
- modular mapping = important feature of DMI design
  $\Longrightarrow$ 2 main conversion mapping layers

Verfaille, Boissinot, Depalle, Wanderley, "*Ssynth*: a real time additive synthesizer with flexible control," ICMC 2006.

Wanderley, Schnell, Rovan, "Escher - modeling and performing composed instruments in real-time," IEEE-SMC'98, pp. 1080–4, 1998.

Mapping layers in *Ssynth*



Mapping layers in *Ssynth*

layer 1. control param. $\longrightarrow$ 'abstract' param. *[Wanderley* et al., *1998]*
$\Longrightarrow$ imitative mapping: models coupling of control param.

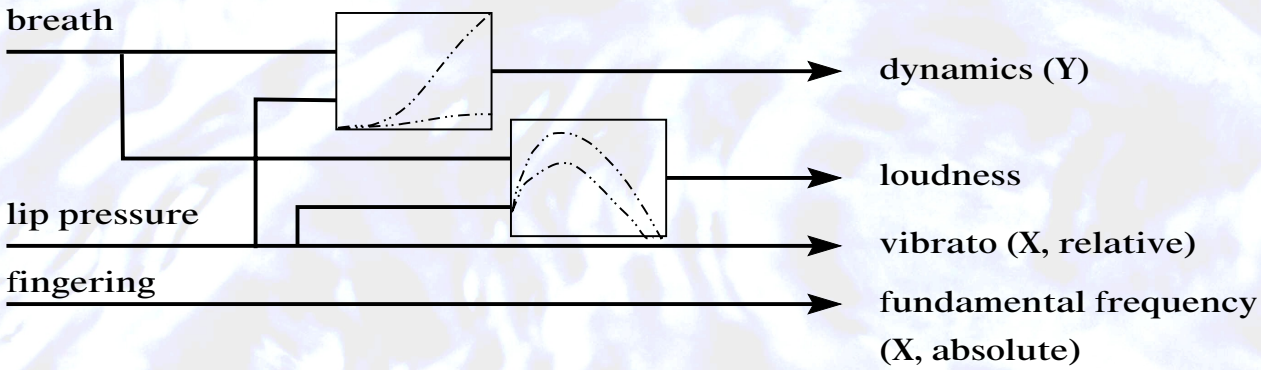


Figure: Second coupling configuration

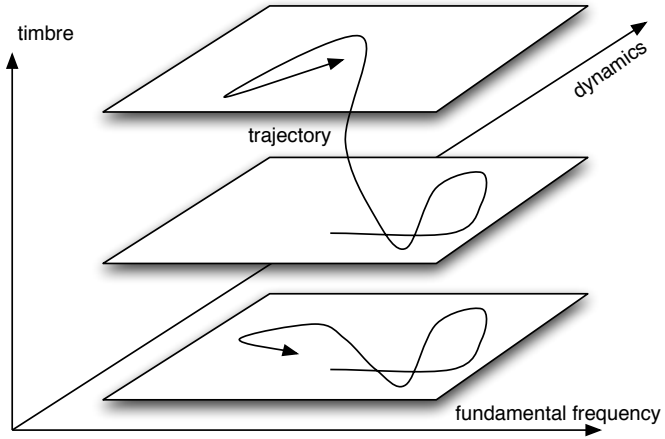
Mapping layers in *Ssynth*

layer 2. abstract param. $\longrightarrow$ additive frame of sound

Database:

- 3D mesh = pitch (7), dynamics (3), instrument (4) + time
- navigation into the additive database by:
  - interpolation (1),
  - extrapolation (1),
  - morphing ($>$ 1 instruments)

  [Verfaille et al.*, 2006]*



Introduction

Description of the synthesis models used

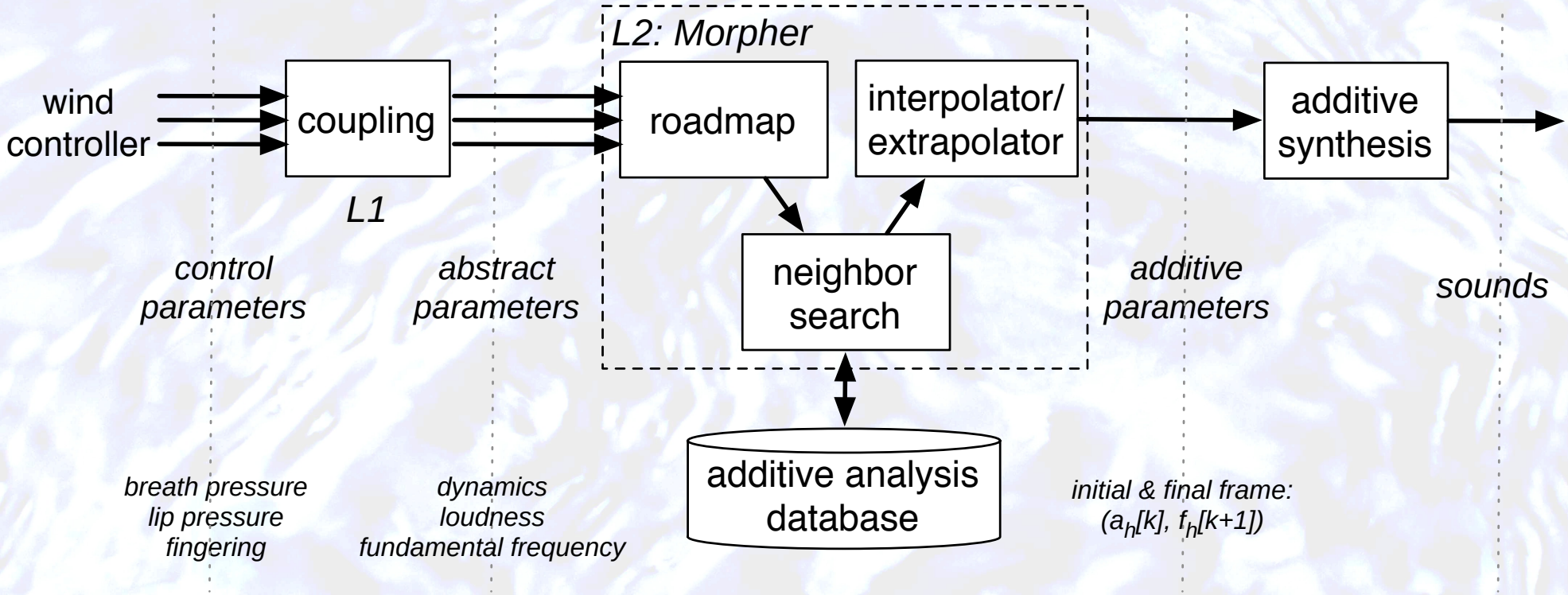
Physically-controlled synthesis

Discussion around an example

Conclusions

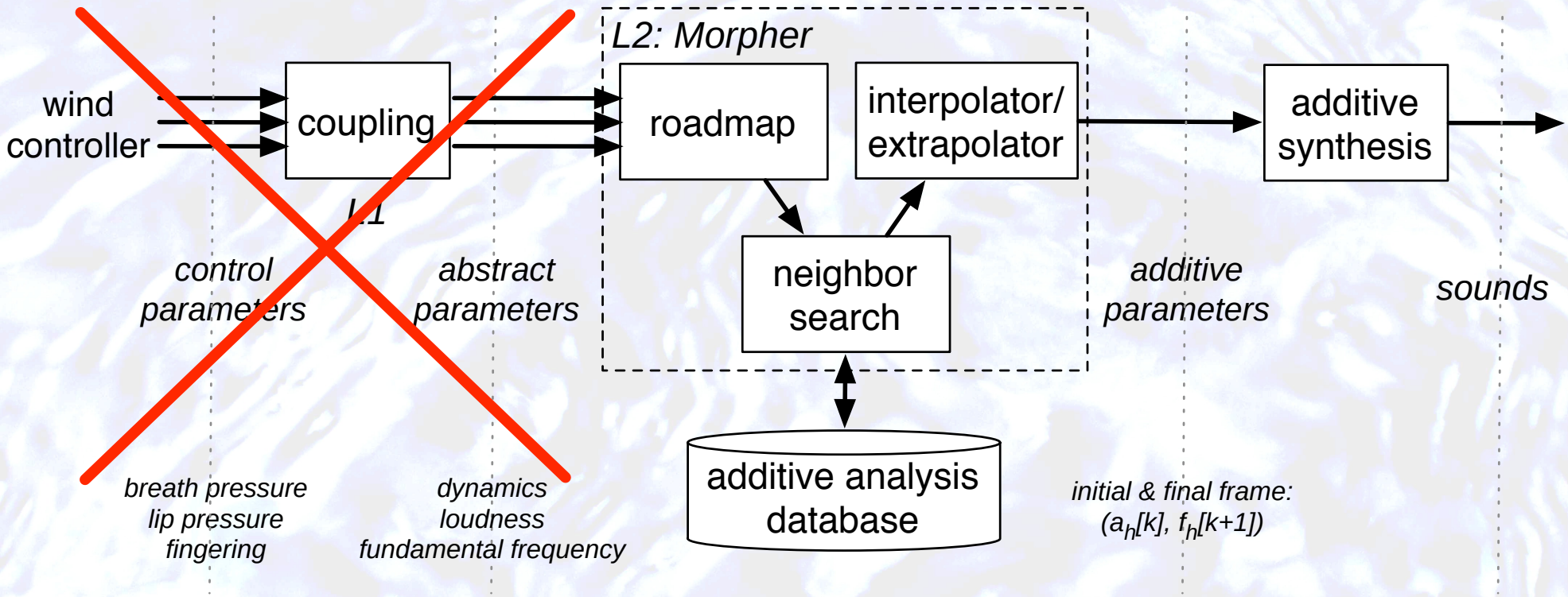
Mapping in physically-controlled synthesis

DIY step 1: from *Ssynth*



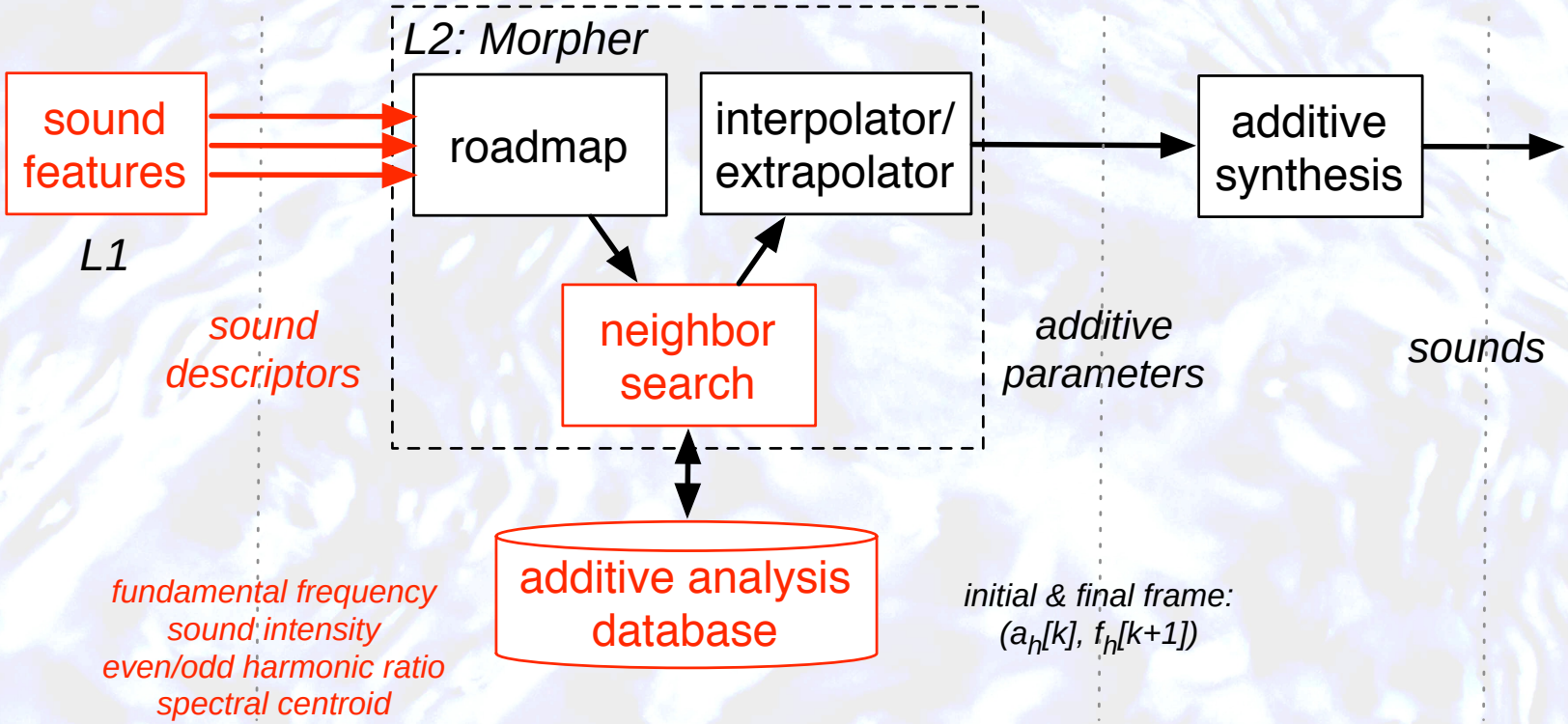
Mapping in physically-controlled synthesis

DIY step 2: remove the first mapping layer (imitative mapping)



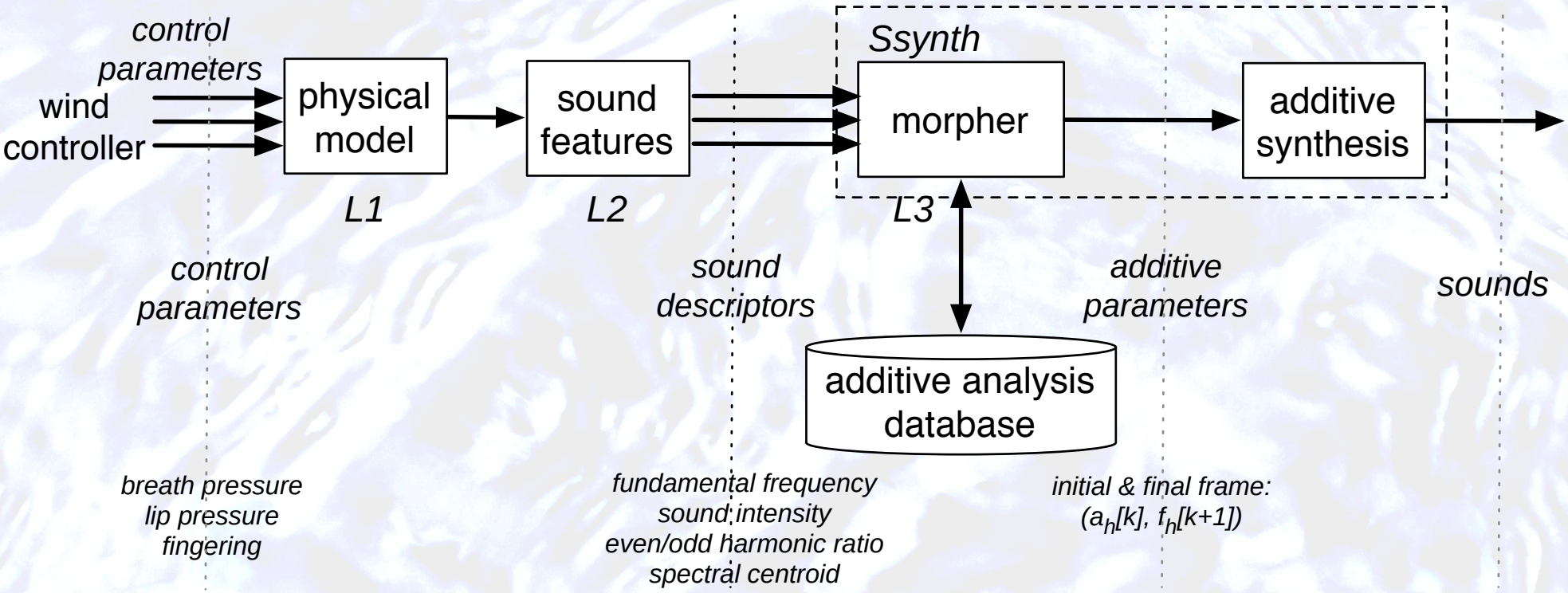
Mapping in physically-controlled synthesis

DIY step 3: extract sound descriptors (index, neighbor search)



Mapping in physically-controlled synthesis

DIY step 4: connect the physical model



Sound descriptors set

Key point: sound descriptors to interface the models

- descriptors linked to perception: pitch, loudness, timbre
  - fundamental frequency: $f_0(k) \in \left[0, \frac{F_s}{2}\right]$,
  - sound intensity: $I(k) \in [0, 1]$,
  - ratio of even/odd harmonics power: $r_{e/o}(k) \in [0, 1]$,
  - spectral centroid: $SC(k) \in \left[0, \frac{F_s}{2}\right]$,
- realtime extraction (goal: improve control ability)
- extracted from the sound
- notorious critical task, can be affected by errors
  $\Longrightarrow$ specific strategies developed

Sound descriptors extraction

Block-by-block analysis driven by $\overline{F_0}$ (WX5):

- physical model: $SC(k)$ / STFT;
  $F_0(k)$, $I(k)$ and $r_{e/o}(k)$ / autocorrelation fct
- additive analysis: using $\{a_i(k), f_i(k)\}$
- spectral & harmonic centroids corrected:
  - Beauchamp: $HC(k, c_0) = \frac{\sum_{h=1}^{H(k)} a_h(k) \cdot f_h(k)}{c_0 + \sum_{h=1}^{H(k)} a_h(k)}$ *[Beauchamp, 1982]*
  - alternative: $\overline{HC}(k, c_0) = \frac{\sum_{h=1}^{H(k)} a_h(k) \cdot f_h(k)}{\max\left(c_0, \sum_{h=1}^{H(k)} a_h(k)\right)}$
- idem for $r_{e/o}(k)$

Beauchamp, “Synthesis by spectral amplitude and ‘brightness’ matching of analyzed musical instrument tones,” JASA (30)6, pp. 396–406, 1982

Introduction

Description of the synthesis models used

Physically-controlled synthesis

Discussion around an example

Conclusions

First results

Sound example with 'Pierre et le loup':

- Physical model controlled by the WX5
- *Ssynth*, database indexed by ($F0$, *dyn*)
- PCS, database indexed by (F0,hc), and by (F0,eob)

So what? Does it sound better that physical modeling/*Ssynth*?

1. database: ***ff***: A3, D4, A4; ***mf***: A3, D4, A4; ***pp***: A3, D4, A4
2. PCS vs. physical model:
   - \+ sounds like *this clarinet*
   - − attacks & brightness changes (due to the database)
3. PCS vs. *Ssynth*:
   - \+ better control ability; mapping is easier to set
   - − descriptor errors, ++ for (F0,eob)$\Longrightarrow$ sound accidents
   - − gain control of attacks ($\sim$ fade in) that sound different

Limitations

List of possible limitations (to our knowledge):

- transients: weakest part of physically-controlled synthesis
  AS for slowly varying sound $\longleftrightarrow$ PM for transients
- sound descriptors quality is critical (not good enough yet)
- additive synthesis: tends to sound like the physical model
  $\Longrightarrow$ non overlapping ranges of sound descriptors!
  (limited additive data set from a real clarinet)
- not a complete physical model of clarinet
  $\Longrightarrow$ dynamic behavior of a given instrument not captured

$\Longrightarrow$ several improvements needed

Improvements

IDs for improving PCS (* with A. Rebillout):

- improve quality of attacks analysis
- use enhanced physical models
  - beating-reed phenomenon
  - consider the network of toneholes
- combine acoustical information from both models
  $\Longrightarrow$ requires satisfactory morphing strategies
- refine / change search strategies of neighbor frames*:
  - search without considering the time unfolding of data
  - replace 2D (f_0, HC) or $(f_0, r_{e/o})$ by 3D $(f_0, HC, r_{e/o})$
  - search strategies depending on attack / sustain

Improvements

IDs for improving PCS (* with A. Rebillout):

- ↗ additive database information*:
  - ↘ granularity: 100 ⟶ 200 Hz; 2 ⟶ 12 notes / octave
  - ↗ representativity: attack types, performers/clarinets
  - ↗ timbre descriptors set (+ time derivatives)

  may result in ↗ neighbor search quality, ↘ performances
- ↘ descriptor computation time to a few samples* (filters):
  - \+ reduces latency (by the block size)
  - − need to compensate for descriptor estimation error

Introduction

Description of the synthesis models used

Physically-controlled synthesis

Discussion around an example

Conclusions

Conclusions

- physically-controlled synthesis introduced as a way to combine two synthesizers
- can be described as a mapping strategy for realtime control
- attempt to keep advantages of both:
  - a physical model (control quality)
  - an additive resynthesizer (sound quality + mapping design)
- interface key point: choice of sound descriptors
- first tests: 2D search and 1-to-1 interface mappings
  $\Longrightarrow$ not fully successful yet (especially transients)!
- future works needed

Thanks!

Any questions, comments, suggestions?

For fans of audio effects: what it does not sound like anymore :-)

Selected bibliography

J. W. Beauchamp, "Synthesis by spectral amplitude and "brightness" matching of analyzed musical instrument tones," *J. Audio Eng. Soc.*, vol. 30, no. 6, pp. 396–406, 1982.

P. Guillemain, J. Kergomard, and T. Voinier, "Real-time synthesis of clarinet-like instruments using digital impedance models," *J. Acoust. Soc. Am*, vol. 118, no. 1, pp. 483–494, 2005.

R. T. Helland, "Synthesis models as a tool for timbre studies," Master's thesis, Norwegian University of Science and Technology, Department of Electronics and Communications, 2004.

J. Kergomard, S. Ollivier, and J. Gilbert, "Calculation of the spectrum of self-sustained oscillators using a variable truncation method: Application to cylindrical reed instruments," *Acustica*, vol. 86, pp. 685–703, 2000.

E. Tellman, L. Haken, and B. Holloway, " Timbre Morphing of Sounds with Unequal Numbers of Features," *J. Audio Eng. Soc.*, vol. 43, no. 9, pp. 678–89, 1995.

V. Verfaille, J. Boissinot, P. Depalle, and M. M. Wanderley, "*Ssynth*: a real time additive synthesizer with flexible control," *Proc. of the Int. Computer Music Conf. (ICMC'06), New Orleans*, 2006.

M. M. Wanderley Ed., *Organised Sound, Special issue on mapping*. Francis & Taylor, 2002, vol. 7, no. 2.

T. A. Wilson and G. S. Beavers, "Operating modes of the clarinet," *J. Acoust. Soc. Am*, vol. 56, pp. 653–658, 1974.